DOCUMENT RESUME

ED 036 405	24	RE 002 386

AUTHOR	TATHAM, SUSAN MASLAND
TITLE	READING COMPREHENSION OF MATERIALS WRITTEN WITH SELECT ORAL LANGUAGE PATTERNS: A STUDY AT GRADES 2 AND 4.
INSTITUTION	WISCONSIN UNIV., MADISON. RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING.
SPONS AGENCY	OFFICE OF EDUCATION (DHEW), WASHINGTON, D.C. BUREAU OF RESEARCH.
REPORT NO	TR-86
BUREAU NO	BR-5-0216
PUB DATE	JUL 69
CONTRACT	OEC-5-10-154
NOTE	142P.
EDRS PRICE	EDRS PRICE MF-$0.75 HC-$7.20
DESCRIPTORS	GRADE 2, GRADE 4, *LANGUAGE PATTERNS, *LANGUAGE RESEARCH, LANGUAGE ROLE, *ORAL EXPRESSION, *READING COMPREHENSION, *READING RESEARCH

ABSTRACT

TO DETERMINE WHETHER OR NOT STUDENTS IN GRADES 2 AND 4 COMPREHENDED MATERIALS WRITTEN WITH PATTERNS THAT APPEAR FREQUENTLY IN THEIR SPEECH BETTER THAN MATERIALS WRITTEN WITH PATTERNS THAT APPEAR INFREQUENTLY, TWO READING COMPREHENSION TESTS WERE DEVISED BY THE INVESTIGATOR. SUBJECTS WERE ALL SECOND- AND FOURTH-GRADE CLASSROOMS FROM TWO SIMILAR SCHOOLS: 163 GRADE-2 STUDENTS (81 GIRLS, 82 BOYS) AND 137 GRADE-4 STUDENTS (69 GIRLS, 68 BOYS). TEST A USED FREQUENTLY OCCURRING PATTERNS FROM THE ORAL LANGUAGE OF SECOND AND FOURTH GRADERS, AND TEST B USED INFREQUENTLY OCCURRING PATTERNS IN THE ORAL LANGUAGE OF STUDENTS FROM THE SAME GRADES. PATTERNS WERE SELECTED FROM STRICKLAND'S STUDY (1962). CHI SQUARE AND T TESTS WERE USED TO ANALYZE THE DATA. RESULTS INDICATED (1) SIGNIFICANTLY MORE GRADE-2 AND -4 STUDENTS OBTAINED HIGHER SCORES ON TEST A THAN ON TEST B (P .001); (2) GRADE-4 STUDENTS PERFORMED SIGNIFICANTLY BETTER THAN GRADE-2 STUDENTS ON BOTH TESTS (P .01); (3) IN GENERAL, THERE WERE NO SIGNIFICANT SEX DIFFERENCES ON EITHER TEST WITHIN OR ACROSS GRADES. TABLES, CHARTS, AND A BIBLIOGRAPHY ARE INCLUDED. (AUTHOR/WB)

RE

86

ED036405

No. 86

READING COMPREHENSION OF MATERIALS WRITTEN WITH SELECT ORAL LANGUAGE PATTERNS: A STUDY AT GRADES 2 AND 4

Report from the Individually Guided Instruction
in Elementary Reading Project

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

RE 002 386

U.S. Office of Education Center No. C-03 Contract OE 5-10-154

BR-5-0216 TR
PA-24
OE/BR

Technical Report No. 86

READING COMPREHENSION OF MATERIALS WRITTEN WITH SELECT ORAL LANGUAGE PATTERNS: A STUDY AT GRADES 2 AND 4

Report from the Individually Guided Instruction
in Elementary Reading Project

By Susan Masland Tatham

Wayne Otto, Professor of Curriculum and Instruction
Chairman of the Examining Committee and Principal Investigator

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

July 1969

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in The University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Launor F. Carter
Senior Vice President on
Technology and Development
System Development Corporation

Francis S. Chase
Professor
Department of Education
University of Chicago

Henry Chauncey
President
Educational Testing Service

Martin Deutsch
Director, Institute for
Developmental Studies
New York Medical College

Jack Edling
Director, Teaching Research
Division
Oregon State System of Higher
Education

Elizabeth Koontz
President
National Education Association

Roderick McPhee
President
Punahou School, Honolulu

G. Wesley Sowards
Director, Elementary Education
Florida State University

Patrick Suppes
Professor
Department of Mathematics
Stanford University

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

UNIVERSITY POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

***James W. Cleary**
Vice Chancellor for Academic
Affairs

Leon D. Epstein
Dean
College of Letters and Science

John Guy Fowlkes
Director
Wisconsin Improvement Program

Robert E. Grinder
Chairman
Department of Educational
Psychology

H. Clifton Hutchins
Chairman
Department of Curriculum and
Instruction

Clauston Jenkins
Assistant Director
Coordinating Committee for
Higher Education

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational
Psychology

Donald J. McCarty
Dean
School of Education

Ira Sharkansky
Associate Professor of Political
Science

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

M. Crawford Young
Associate Dec
The Graduate School

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of
Sociology

Max R. Goodson
Professor of Educational Policy
Studies

Russell J. Hosler
Professor of Curriculum and
Instruction and of Business

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational
Psychology

Wayne Otto
Professor of Curriculum and
Instruction (Reading)

Robert G. Petzold
Associate Dean of the School
of Education
Professor of Curriculum and
Instruction and of Music

Richard L. Venezky
Assistant Professor of English
and of Computer Sciences

FACULTY OF PRINCIPAL INVESTIGATORS

Ronald R. Allen
Associate Professor of Speech
and of Curriculum and
Instruction

Vernon L. Allen
Associate Professor of Psychology
(On leave 1968-69)

Nathan S. Blount
Associate Professor of English
and of Curriculum and
Instruction

Robert C. Calfee
Associate Professor of Psychology

Robert E. Davidson
Assistant Professor of
Educational Psychology

Gary A. Davis
Associate Professor of
Educational Psychology

M. Vere DeVault
Professor of Curriculum and
Instruction (Mathematics)

Frank H. Farley
Assistant Professor of
Educational Psychology

John Guy Fowlkes (Advisor)
Professor of Educational
Administration
Director of the Wisconsin
Improvement Program

Lester S. Golub
Lecturer in Curriculum and
Instruction and in English

Max R. Goodson
Professor of Educational Policy
Studies

Warren O. Hagstrom
Professor of Sociology

John G. Harvey
Associate Professor of
Mathematics and Curriculum
and Instruction

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational
Psychology

Burton W. Kreitlow
Professor of Educational Policy
Studies and of Agricultural
and Extension Education

Richard G. Morrow
Assistant Professor of
Educational Administration

Wayne Otto
Professor of Curriculum and
Instruction (Reading)

Milton O. Pella
Professor of Curriculum and
Instruction (Science)

Thomas A. Romberg
Assistant Professor of
Mathematics and of
Curriculum and Instruction

Richard L. Venezky
Assistant Professor of English
and of Computer Sciences

MANAGEMENT COUNCIL

***Herbert J. Klausmeier**
Director, R & D Center
Acting Director, Program 1

Mary R. Quilling
Director
Technical Section

Thomas A. Romberg
Director
Programs 2 and 3

James E. Walter
Director
Dissemination Section

Dan G. Woolpert
Director
Operations and Business

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Individually Guided Instruction in Elementary Reading Project in Program 2. General objectives of the Program are to establish rationale and strategy for developing instructional systems, to identify sequences of concepts and cognitive skills, to identify or develop instructional materials associated with the concepts and cognitive skills, and to generate new knowledge about instructional procedures. Contributing to these Program objectives, the Reading Project staff, in cooperation with area teachers, prepared a scope and sequence statement of reading skills for the elementary school as a first step in the development of an instructional program. From this outline, assessment procedures and group placement tests have been developed, and existing instructional materials have been keyed to the outline. Research is conducted to refine the program and to generate new knowledge which will be incorporated into the system.

CONTENTS

List of Tables and Figures vii

Abstract ix

I Background of the Study and Related Literature 1
  - Background of the Study 1
  - Purpose of the Study 4
  - Definition of Terms 5
  - Review of the Literature 6
    - Studies of Elementary School Children's Oral Language 7
    - Experimental Studies Using Written Material Based on an Analysis of Children's Oral Language 30
    - Summary 34

II Method 35
  - Overview of the Study 35
  - Development of the Reading Comprehension Tests 36
    - Selection of the Patterns 37
    - Determination of the Patterns to be Used 41
    - Control of Select Variables 44
      - Vocabulary Control 44
      - Content Control 45
      - Grammatical Complexity Control 47
    - Measurement of Reading Comprehension 51
    - Final Selection of Test Items 56
    - Composition of the Final Tests 60
  - Subjects 61
  - Data Collection 64
  - Data Analysis 65
  - Summary 68

III Results and Discussion 69
  - Results of the Tests of Hypotheses 70
  - Results Related to Reading Comprehension of Select Oral Language Patterns 70
    - Each Grade as a Separate Group 72

Each Sex as a Separate Group Within Grades 2 and 4 73
Results Related to Grade Level 74
Results Related to Sex 75
Summary of Results 84
Discussion of the Results 85
Results Related to Reading Comprehension of Select Oral Language Patterns 85
Results Related to Grade Level 88
Results Related to Sex 91
Reliability of the Reading Comprehension Tests 92
Summary 96
Validity of the Reading Comprehension Tests 96
Construct Validity 97
Content Validity 101
Summary 102
Summary 103

IV Summary, Conclusions, and Implications 104
Background and Purpose 104
Related Research 105
Method 106
Results 109
Limitations 110
Conclusions 111
Implications 111
Implications for Beginning Reading Instruction 112
Implications for Teacher Training 112
Implications for Classroom Materials 114
Implications for Measurement of Language-Related Abilities 114
Implications for Research 115

Appendix A
The Twenty-five Most Frequent Oral Language Patterns of Second and Fourth Graders as Identified by Strickland (1962) 117

Appendix E
Frequency Distributions of Scores on Test A and Test B Within and Across Grades 2 and 4 124

Bibliography 137

Appendices B, C, D, and F are not printed herein. See original manuscript in the University of Wisconsin Memorial Library.

LIST OF TABLES AND FIGURES

Table		Page
1	The Rank Order and Frequency of Oral Language Patterns Selected for Test A	40
2	The Rank Order and Frequency of Oral Language Patterns Selected for Test B	41
3	Index of Proportions of Frequent Oral Language Patterns in Material Written for Second and Fourth Graders	43
4	Index of Proportions of Infrequent Oral Language Patterns in Material Written for Second and Fourth Graders	43
5	"Hard" Words Used in Test A and Test B	46
6	Ratio of the Number of Different Words (Type) to the Total Number of Words (Token) in Test A and Test B	46
7	Analysis of Sentence Content in Test A and Test B	48
8	Analysis of Picture Content in Test A and Test B	48
9	Total Index Ratings (IR) Given to Slots in the Oral Language Patterns of Test A	51
10	Total Index Ratings (IR) Given to Slots in the Oral Language Patterns of Test B	52
11	Summary of Total Index Ratings (IR) Given to Slots in Oral Language Patterns Common to Test A and Test B	53
12	The Mean, Range, and Standard Deviation of IQ Scores for All Subjects and for Subjects Grouped by Grade, Sex, and Sex within Grade	62
13	Proportion of Subjects in Three Levels Defined by the Duncan Socio-economic Index Scale	63
14	Chi Square Analyses of Reading Comprehension Test Scores	71

Table Page

15 Mean, Range, Variance, and Standard Deviation for All 300 Subjects on Test A and Test B 72

16 Results of the t test of Mean Scores for Second and Fourth Graders 76

17 A Comparison of the Mean, Range, Variance, and Standard Deviation of Scores for Grades 2 and 4 77

18 Results of the t test of Mean Scores for Boys and Girls Across Grade Levels 78

19 Results of the t test of Mean Scores for Boys and Girls in Second Grade 80

20 Results of the t test of Mean Scores for Boys and Girls in Fourth Grade 82

21 Mean, Range, Variance, and Standard Deviation of Scores on Test A and Test B for Girls and Boys Within and Across Grades Two and Four 83

22 Summary of the Results of the Tests of Hypotheses 86

23 The Number of Poor, Average, and Good Readers Among Fourth Grade Girls and Boys as Determined by Scores from the Paragraph Meaning Section of the Stanford Achievement Test 88

24 Hoyt Reliability Coefficients Within and Across Grades for Test A and Test B 94

25 Partial Regression Coefficients of Test A and Test B 100

26 Partial Correlations of Test A and Test B with the Paragraph Meaning Section of the Stanford Achievement Test 100

27 Validity Coefficient of Correlation between Test A, Test B, the Combined Scores, and the Paragraph Meaning Section of the Stanford Achievement Test 101

Figure

1 Sample Item from Test B 55

2 A Poor Distractor in the Pilot Study from Test B (the Middle Picture) 58

3 The Revised Distractor from Test B (the Middle Picture) 59

ABSTRACT

To determine whether or not second and fourth graders comprehend material written with oral language patterns that appear frequently in their speech better than material written with patterns that appear infrequently, two reading comprehension tests were devised by the investigator and given to all subjects. Subjects comprised all second and fourth graders in two similar schools; 163 second graders (81 girls, 82 boys) and 137 fourth graders (69 girls, 68 boys) were used. Test A used patterns that appear frequently in second and fourth graders' oral language (e.g., noun-verb-object), and Test B used patterns that appear infrequently in the oral language of the same grades (e.g., noun-verb-indirect object-direct object); patterns were selected from Strickland's study (1962). Chi square analysis was used to determine differential effects of frequent and infrequent oral language patterns on reading comprehension; t tests were used to determine differences between grade levels and sex. Results included the following: (1) significantly more second and fourth graders obtained higher scores on Test A than Test B ($p < .001$); (2) fourth graders performed significantly better than second graders on both tests ($p < .01$); (3) in general, there were no significant sex differences on either test within or across grades.

I

BACKGROUND OF THE STUDY AND RELATED LITERATURE

Background of the Study

This study concerns the relationship between reading comprehension and material written with language patterns that appear in the oral language of second and fourth graders. Prior to 1959, conducting a study of this kind would have been a great problem, as there was no objective and reliable method for analyzing the speech of large numbers of children. Early attempts to study children's language were likely to be diaries of single children. Information collected in this manner was obviously not a basis for broad generalizations.

In 1959, however, a group of linguists and others devised a practical and useful method for analyzing children's speech. This scheme has been used in at least four major studies which provide information about the linguistic structure of children's oral language (Strickland, 1962; Hocker, 1963; Loban, 1963; Riling, 1965). In two of these studies, the newly acquired findings about children's oral language ability were related to children's reading ability. This line of inquiry could very well have been prompted by the fact that elementary school teachers generally use a variety of methods to provide oral language experiences as a prerequisite to beginning reading instruction; yet very little is known about the nature of the assumed relationship between these two skills. The researchers found that the more mature users of the language, generally defined as those children who manipulated the basic patterns with greater flexibility, were also the better readers (Strickland, 1962; Loban, 1963). The findings suggest that facility with spoken language is related to successful comprehension of written language patterns.

Since children's abilities to produce and comprehend spoken language are not evident to the same degree at a particular time (Vygotsky, 1962), it is likely that school children's ability to produce oral language patterns is, at some stage, also widely different from their ability to read these patterns. A reason for this discrepancy is provided by linguists like Lefevre (1964), who have attempted to show that reading is not, contrary to popular definition, just "speech written down." Intonation qualities of speech that provide valuable meaning clues are absent in "speech written down"; the reader must learn conventions that are unique to written language such as capital letters and paragraphing. Whereas the process of learning to speak is apparently effortless and unconscious, the process of learning to read is obviously not effortless for many children. The enormous variety of teaching methods and materials and the concomitant controversy over their use attest to this statement.

Several linguists and educators have discussed some of the unnecessary obstacles that the young reader must overcome before he learns how spoken language is represented in writing. Stevens (1965) criticized the use of "unnatural language" and "absurdly unnatural sentence rhythm" in most basal readers. Amsden (1964) suggested that children may fail in reading because of the use of teaching materials which are poorly designed in terms of their oral language patterns; he concluded that they may be reading a syntax that is too hard for them. Goodman (1968) has observed children inserting words as they read in an attempt to relate their customary way of speaking to the written syntax of basal readers. Strickland (1962) found that there appeared to be no scheme for the development of control over sentence structures in the texts she analyzed; language patterns seemed to be introduced at random. Riling (1965) found that sixth grade textbooks do not use language in a way which is much superior structurally to the language of fourth grade textbooks.

From observations like these, it is apparent that children's reading comprehension is often obstructed by materials that unnecessarily obscure the relationship between spoken and written language—a relationship which, according to linguists, is the crux of reading behavior (Carroll, 1964). Using materials structured more like the way children speak is a logical alternative to faulty comprehension caused by "unnatural language." Materials constructed this way would make use of an important principle from learning theory, which defines efficient learning as that which "allows the learner to start with those learnings he brought with him to the task and upon these gradually build new responses which take him nearer and nearer to the desired terminal behavior [Carroll, 1964, pp. 348–9]." In this case, the "learnings" a child brought with him would be his way of structuring what he has to say. The "new responses" would be his increasing ability to read language patterns that differ more and more from the ones he uses in his oral language.

One study has focused on the possibility of improving children's reading comprehension by using materials structured more like the way they speak (Ruddell, 1963). Its major finding was that fourth graders comprehended material written with patterns that appear frequently in children's oral language significantly better than material written with patterns that appear infrequently in the oral language of children. More research is needed to confirm these results for different groups of children and at different grade levels. Determining whether or not certain language patterns used by children are easy or difficult for them to comprehend in written form could bring a relevant but heretofore neglected characteristic of the reader—his oral language—to the concept of readability. Research in this area could thus begin to define the nature of the relationship between oral language facility and reading ability.

Purpose of the Study

The present study was undertaken to investigate further the relationship between reading comprehension and material written with select oral language patterns at second and fourth grades. The relation of sex to reading comprehension of these patterns was of additional interest, as it was hoped that this study could contribute to the dialogue between those who have found significant sex differences in language abilities of children and those who have not. Comparing the performance of boys and girls at two grades was intended to encourage the emergence of any developmental levels with respect to children's ability to comprehend diverse kinds of language patterns. Fourth graders were included as subjects so that a comparison with the findings of the other study in this area would be possible. Second graders were added to provide some information about children who are considerably less experienced in relating spoken and written language.

The study was designed to answer three basic questions:

Q1: Do significantly more second and fourth graders comprehend material written with frequent oral language patterns better than material written with infrequent oral language patterns?

Q2: Do fourth graders comprehend material written with frequent and infrequent oral language patterns significantly better than second graders?

Q3: Do girls comprehend material written with frequent and infrequent oral language patterns significantly better than boys?

Definition of Terms

Several terms are used frequently throughout the following chapters. Since definitions are often confined to the context of the study, the terms are listed here for the reader's convenience.

Language pattern.—A language pattern is a unit in oral or written language that indicates the sequence of slots and movables. Examples are the noun–verb–object pattern, coded as 1 2 4 (She jumps the log), and the noun–verb–adverb-of-manner pattern, coded as 1 2 M_2 (She jumps quickly).

Sentence.—A sentence is a single language pattern, or several patterns joined by a conjunction, that begins with a capital letter and ends with a period.

Phonological unit.— Used in the analysis of children's oral language, this unit is determined by the inflection, stress, and pause of the speaker; oral language is segmented into phonological units according to the sound-system of the English language (Loban, 1963). This term has been applied inaccurately as a synonym for "sentence."

T-unit.— Popularized by Hunt (1966) in his work with analyses of children's written language, this term is defined as a main clause with any attached subordinate clauses. The T-unit is considered a more accurate and thus more meaningful measure than the phonological unit.

Slot.— A slot is an immovable unit in a pattern. In this study, these units are designated in the following manner: 1 = subject; 2 = verb; 2b = passive verb, verb of the to be class, or copulative verb; 3 = indirect object; 4 = direct object; 5 = predicate nominative.

Movable.— A movable is a unit of a pattern that generally does not have a fixed position. The movables in this study have been categorized into three subgroups: M_1 = an adverb of place (e.g., in the house;

here); M_2 = an adverb of manner (e.g., quickly; on foot); M_3 = an adverb of time (e.g., in winter; then).

T marker.— A T marker used before a pattern indicates that the pattern is preceded by another pattern. A T marker used after a pattern indicates that the pattern is followed by another pattern. Joining two patterns with T markers, 1 2 4 +T and T 1 2 4, would result in 1 2 4 + 1 2 4, illustrated by the sentence, "The dog sees the cat and the cat sees the dog."

Frequent oral language pattern.— For purposes of the study, this type of pattern appeared among the ten most frequent patterns in the oral language of second and fourth graders as identified by Strickland (1962).

Infrequent oral language pattern.— An infrequent pattern was one that ranked lower than tenth in frequency in the oral language of second and fourth graders as identified by Strickland (1962).

Slot filler.—A slot filler comprises the single words, adjectives, phrases, or clauses with which a slot may be filled.

Stanford subtest.— This term refers to the Paragraph Meaning section of the Stanford Achievement Test taken by the subjects.

Review of the Literature

The literature reviewed in this chapter is divided into two parts. The first section, studies of elementary school children's oral language, is limited to investigations that have been based on a broad analysis of oral language obtained from relatively large samples of elementary school children. Studies of children's language development have been reviewed in the Encyclopedia of Educational Research (Carroll, 1960), the April issue of the Review of Educational Research (Carroll, 1958;

Kjeldergaard, 1961; and Carroll, 1964), and a more recent publication by O'Donnell, Griffin, and Norris (1967). The intent in this chapter is to focus on studies that have used a more sophisticated and reliable method of analyzing children's oral language than existed before 1959. The number of studies is consequently small—five. The second section, experimental studies using written material based on an analysis of children's oral language, is confined to a discussion of a single study, which is believed to be the only one in this area at present.

Studies of Elementary School Children's Oral Language

In this section, the five broad analyses of children's oral language reported since 1959 are reviewed. The first study of its kind was published by Ruth Strickland in 1962. She undertook to study oral language samples of children from first through sixth grades with these purposes: (1) to analyze the structure of the oral language, (2) to compare the structure of the oral language with that found in children's texts, and (3) to determine whether or not characteristics of the oral language structure were related to the quality of a child's reading ability.

The sample comprised 575 children who were randomly selected from the 16 public schools in Bloomington, Indiana. There were 100 subjects in each of Grades 1—3, 5, and 6, and 75 subjects in the fourth grade. The variables of age, sex, intelligence, and socioeconomic status were studied at each grade level, and the variable of parental education was considered in most of the grades. Some additional analyses were made at the sixth-grade level; children's oral language was compared to their silent reading comprehension, their listening comprehension, and their oral reading interpretation. The methodology and findings for each of the three major, relevant concerns of the study will be discussed separately below.

All of the oral language samples were obtained by having children come to a testing room at the school in groups of twos and threes. The pupils sat at a table which had several familiar storybook characters on it. The examiner encouraged them to talk freely about anything that interested them; the figures and the examiner served to stimulate the conversation.

The conversations were taped until each child was "talking naturally." The language samples were then transcribed, and 25 sentences or phonological units[1] from each child's sample were analyzed for these characteristics: (1) syntactic structure of sentences, (2) frequency of occurrence of certain language patterns, (3) amount and kinds of subordination, (4) length of sentences, and (5) flow of language. The relationship of select language variables to age, sex, and intelligence levels of the children was then ascertained.

---

[1] In Strickland's study, the term sentence meant phonological unit, which she defined as a "unit of speech ending with a distinct falling intonation which signals a terminal point [p. 16]." The unreliability of this unit, however, has been demonstrated by researchers like Kean (1967) and Riling (1965), who attempted to explain how she distinguished between long phonological units joined by and and a series of short phonological units also joined by and. Her decision was based on (1) the pitch which marked the beginning and end of the unit, (2) the length of the pause before and, (3) the location of the falling intonation, and (4) the pronunciation of and. Some of these criteria are surely difficult to judge. A more reliable unit, the T-unit, has been developed and tested by Hunt (1965) in his studies of children's structures in written language. This unit consists of "one main clause with all the subordinate clauses attached to it [p. 20]." The use of this measure should provide more reliable results for future studies of children's oral language.

A singular feature of this study at the time of publication was the scheme of analysis used on the 14,375 phonological units that comprised the language sample. The scheme was the product of linguists who met for this purpose in 1959 at Indiana University.[2] Two distinct kinds of analyses were outlined: Level I, in which sentences are divided into fixed slots (the immovable units in a language pattern such as the subject and verb) and movables (the unit of a pattern that generally does not have a fixed position), and Level II, in which sentences are analyzed for the type of subordinations used in the slots and movables.[3]

Since Strickland's study was primarily a description of the oral language patterns, the findings were presented in tables which summarized the frequency with which certain patterns appeared for select groups of children. Tests of significance to determine trends across the grades were not made; it is usually up to the reader to draw his own conclusions about the meaning of figures in the tables. The only statistic applied to the data was chi square, which was used to study the relationship between certain language variables and background data of the subjects. In sum, Strickland's study was not a statistical analysis, but rather a descriptive one. The findings should be considered in this light.

Some of the major findings of her analysis of children's oral language are summarized briefly below:

1. Data for the word-length of phonological units were presented for Grades 1,4, and 6 in two tables. In one, the length of phonological units

---

[2] The participants were John Carroll, W. Nelson Francis, David Reed, Fred Householder, Harold Whitehall, Walter Loban, Virginia Mini, Mansur Ekhtiar, Eldonna Everetts, and Ruth G. Strickland.

[3] For a full description of these levels, the reader is referred to pp. 18–22 of Strickland's study (1962).

was summarized in terms of mean and range by grade and sex; differences between means were not tested for significance. In the second table, the mean length of phonological units for the highest and lowest quarters of pupils in Grades 2 and 4 was summarized according to age, intelligence, mental age, occupational status, and father's and mother's education. As no tests of significance between means were made, the conclusions were stated in a general way: (1) Sentences varied considerably in length, the range at first grade being from 1-76 words, for example. Differences between means and variables of the subjects' background at Grades 4 and 6 were small but always in favor of the high group. (3) Length varied more within a grade than from grade to grade.

2. With respect to the occurrence of language patterns, Strickland found that the number ranged from 658 in the sample for first grade to 1,041 for the sixth grade. Although a wide range of language patterns was used at all grade levels, some appeared much more frequently than others at all the grade levels. Strickland suggested that the relatively few language patterns are the "basic building blocks" of children's language. A table of the rank order and frequency of the structural patterns most commonly used by children was presented for Grades 1, 2, 4, and 6. Of the 5 highest ranking patterns that were common to each of these grades, none contained a movable element.

3. Results of the Level II analysis were presented in a table which indicated, for Grades 1, 2, 4, and 6, the per cent of use of each type of subordinate element and the per cent of the sample at each grade level that used the element. It is clear from the table that children at all grade levels expanded and elaborated their sentences through the use of movables and elements of subordination. Differences between grade levels were not tested for significance, or even discussed, so the presence of

developmental levels was not readily apparent if, indeed, these levels exist in the data.

4. By means of chi square tests, the relation between the use of movables and subordination and select variables was ascertained. Differences significant at the .05 level or better were found between the use of subordination and movables by subjects in Grades 1, 4, and 6 who differed in verbal intelligence, mental age, occupational status, and parents' education.

In sum, these generalizations can be derived from Strickland's analysis of children's oral language:

1. Children at all grade levels use a wide range of language patterns.
2. Certain patterns which children use with great frequency appear to be basic to their language.
3. Children at all grade levels can expand and elaborate their sentences through the use of movables and elements of subordination.
4. There are significant differences between the use of movables and patterns of subordination and variables of intelligence, mental age, occupational status, and parental education.

A second major focus of the study was the comparison of children's language with the language of their textbooks. For this purpose, Strickland selected four sets of readers for analysis. Each preprimer and eight pages from the other books comprising each set were studied to find out which of the patterns that were used most frequently by children in Grades 1, 2, 4, and 6 appeared in the texts for those grades. Two points need to be made to place the findings in perspective. First, certain patterns were selected from children's language and compared with patterns in the texts. No attempt was made to report those patterns in the text that did not coincide with these selected oral language patterns. Second, as Strickland was careful to state, this analysis was not meant to be a frequency tabulation of the patterns in the books. The intent was merely to state whether or not certain patterns had appeared at all on the sample pages.

The findings were presented in a table which shows the rank order of the selected frequent oral language patterns for each grade level and whether or not the pattern appeared in the texts analyzed for that grade.

Within the limitations of the sampling procedures, these conclusions were drawn:

1. The patterns which appeared in the sample were not the same from book to book either within or across any of the series.

2. The patterns appeared to be introduced somewhat at random in the texts.

3. There appeared to be no scheme for the development of control over sentence structures in the texts that parallels the generally accepted scheme for the development of control over vocabulary.

A third major concern of the study was to determine the relationship between oral language and certain reading abilities. A study of the relationship between oral language and reading achievement was made for Grades 2 and 6 only. The Paragraph Meaning test from the Gates Primary Reading Tests (Type AWR, Form 3) was used to provide information about the reading comprehension of the 100 second graders. The scores obtained from this test were translated into reading age. Subjects were grouped into one of five categories of reading age, and a tabulation of the frequency and average use of each of the most frequent language patterns was made. Other comparisons between the groups of children were made in terms of sentence length, the use of movables, and the use of subordination. The results were presented in several tables, and a few case studies of representative children were added.

Pertinent findings included the following:

1. Children in the lowest category of reading age were characterized by greater use of the basic patterns than children in other categories; the basic patterns contained no movables.

2. Children whose reading scores placed them above the grade norm tended to use longer sentences.

3. When the children in the lowest and highest categories were compared, there were no significant differences in their use of movables or their use of subordination patterns according to the chi square analysis performed.

At the sixth-grade level, three measures of comprehension were obtained: (1) silent reading comprehension ascertained by the Paragraph Meaning test of the Stanford Achievement Test (Intermediate Level, Form L), (2) oral reading interpretation defined by scores from the oral part of Section IV, Form A, Grades 1 to 8, of the Diagnostic Reading Tests (This test was administered individually and the oral reading was taped.), and (3) listening comprehension measured by a group test in which children listened to taped selections and answered written questions afterward.

The sixth graders were grouped on the basis of their grade placement scores on the silent reading comprehension test. The frequency with which select language patterns appeared for each of the resulting five groups was presented in a table. Similar tables were presented for subjects grouped by oral reading interpretation, which was rated along a five-point scale devised by Strickland for the study, and by listening comprehension, which was measured by the number of correct responses to the listening test. Comparisons of mean sentence length, the use of movables, and the use of subordination were made for those subjects who were rated high or low on any of the three comprehension tasks. Results of the chi square analyses were presented for the last two comparisons between groups.

The findings can be summarized as follows:

1. There was a relationship between the structure of children's oral language and silent reading comprehension. Pupils who ranked high in silent reading made greater use of movables and elements of subordination than children who ranked low. This difference between groups was significant at the .02 level. Pupils who ranked high also had a greater mean sentence length than children in the low group, but this difference was not tested statistically.

2. There was a relationship between oral reading interpretation and the structure of children's oral language. A difference significant at the .01 level was found between the high and low groups in oral reading interpretation and the use of movables and subordinate elements.

3. There was a relationship between listening comprehension and the structure of children's oral language. Children who were rated high in listening comprehension had a greater average use of movables and subordinate elements and a longer mean sentence length than children ranked low, but these differences were not statistically significant.

4. Children ranked high in silent reading, oral reading, and listening used longer sentences than children ranked low in these areas. The differences, however, amounted to 2.31 words at most.

Strickland's work has been criticised for its restricted child population, the attempt to compare informal conversation with edited reading material (Burrows, 1964), and its failure to differentiate details of subordination at the Level II analyses performed (Mayer, 1964). In general, however, this study has been marked as a significant contribution to descriptions of elementary school children's oral language.

Evidence of the effect of Strickland's work (1962) on subsequent research is clear. In the years immediately after its publication, two studies emerged which distinctly followed her methodology. The first was a master's thesis by Mary Elsa Hocker (1963). Her study of children's oral language was undertaken with the assumption that "the more nearly the printed material resembles the child's oral language patterns, the easier it should be for him to anticipate meaning in the sentences he is beginning to read [p. 1]." Being interested in the beginning stages of reading, she focused on about 40 first graders from two schools at an Air Force Base in Arizona. Her methodology was unique in that language samples were collected in a variety of situations. Birthday parties, Sunday school, general play, and classroom show-and-tell provided some of the settings in which she worked—a distinct break away from the more structured situation used by Strickland (1962) and later by Loban (1963) and Riling (1965). Consecutive utterances of the subjects were recorded by the examiner who used the stenographic method and tape recorder, the latter when possible.

A total of 2500 phonological units were collected and analyzed for (1) language pattern, (2) frequency of occurrence of these patterns, (3) length of sentences, (4) the kind of vocabulary used, and (5) interests. The first three concerns were also important to Strickland's work, while the last two were added by Hocker. As in Strickland's study, a primary concern was the segmentation of phonological units into distinct language patterns. For this purpose, then, Level I of the linguistic scheme previously discussed was used. It will be remembered that this level provides for segmenting any phonological unit into its slots (the fixed elements of a sentence) and movables (the sentence elements that generally do not have a fixed position).

The findings were presented in a series of tables, one of which listed each pattern and the frequency with which it occurred. Statistics were not applied to the data so that, as in the greater part of Strickland's work, the results were broad descriptions of general trends rather than statistical evidence of trends.

Among those findings that are most pertinent to the present study are the following:

1. The mean sentence length of the language sample was 4.86 words, which contrasts with 6.6 for the six-year-olds of Templin's study (1957) and 11.04 and 10.70, respectively, for the first grade boys and girls of Strickland's study (1962).

2. Of the five basic sentence patterns that occurred in the sample, three of them were dominant and formed the syntactic structure for numerous variations in the children's oral language: 1 2 4 (noun—verb—object), 1 2b 5 (noun—to be verb—predicate nominative), and 1 2 (noun—verb). Although the first two patterns were among the most frequent in Strickland's sample at all grade levels (1962), the 1 2 pattern was generally considerably less frequent.

3. The trend of the pattern frequency moved from the simple to the more complex in all cases. This finding supports Strickland's more detailed observations; the most frequent patterns were those without movables.

Though more humble than Strickland's work, Hocker's study contributes several important points that have not been developed to any extent beyond theory by other researchers in the area. First is her conclusion that the kind of oral language a child uses varies with the situation. Hocker apparently made this statement after considering the variance of sentence length of her subjects from situation to situation and the contrast between a child's verbal encounter with an adult and with another child. Unfortunately, a thorough investigation of the noted differences was not made. It is logical to assume that slight, or even significant, differences in the findings of studies of children's oral language need to be considered with respect to (1) the kind of stimuli used in obtaining the language sample, and (2) the situation in which the oral language was sampled.

A second conclusion grew from observations of differences in the language depending on where the language was used. Hocker suggested that because the language patterns used are influenced by the situation in which the child is presently operating, it would "be a mistake" to provide him with a description of a school situation in the language patterns of the playground. Hocker did not differentiate between playground language patterns and schoolroom patterns explicitly enough to support this generalization. The idea, however, should be considered in future research in which the relationship between children's oral language and written material is studied.

Thus, although Hocker used the same scheme as Strickland for analyzing the oral language, unlike Strickland she varied the context in which the language samples were collected, and she limited her investigation to first graders. Perhaps the major contribution of this work is the attempt to begin answering a question posed in Strickland's study: Would children be "aided or hindered by the use of sentences in their books more like the sentences they use in their speech [1962, p. 106]?"

Riling's study (1965) followed Strickland's methodology closely. The influence of the earlier work was made explicit in Riling's statement of purpose:

> This study is designed to analyze the structure of children's written and spoken language in the fourth and sixth grades; to ascertain the influence of certain variables such as intelligence, sex, socioeconomic status of the family, education of the parents, and ethnic origin on the form of the language of these children; and to compare the results of this study

> with those obtained in a similar study conducted by Ruth G. Strickland of Indiana University [p. 3].

Riling was also interested in comparing the patterns which children use at these grade levels with those found in reading textbooks designed for these grades.

A total of 300 children were drawn from six communities within a radius of 70 miles of Southeast Oklahoma State College; the sample comprised 114 Caucasian and 110 Negro children for Grade 4 and 110 Caucasian children for Grade 6. The composition of the sample contrasts with Strickland's (1962) undifferentiated group of 75 subjects in Grade 4 and 100 in Grade 6. The procedure for obtaining the language samples differed from those of both Strickland and Hocker (1963). Perhaps to avoid criticism in comparing oral and textbook language, Riling tried to keep the narrative style of the texts that were to be analyzed constant with the narrative oral and written language of the children. Oral language was stimulated by a picture that was placed before the child, who came into the testing room alone and told his story about the picture; the written language sample was stimulated by a different picture that was used by the investigator in a regular classroom setting. The children discussed the picture and then wrote a story about it.

The process of taping and analyzing the language was similar to Strickland's method. Twenty-five "independent verbalizations of oral . . . and written language" (the number also selected by Strickland) were the maximum amount of language analyzed for each child. For her comparison of children's language with the language of textbooks, Riling designated pages 1, 10, 20, and 50 from the front of each book and four pages in the same relative order from the back as the sample. She followed this procedure for six commonly used textbooks for each grade. The sentences were analyzed by the two-level scheme previously discussed—Level I showing the arrangement of basic grammatical elements and Level II showing elaboration of these elements. To obtain a measure of reading comprehension, the Stanford Achievement Test (also used in Strickland's study) was given to the subjects.

Among the conclusions that are relevant to the present study are the following:

1. Children with higher intelligence, more highly educated parents, and higher family socioeconomic status do not show any significant superiority in the use of the simplest, most basic language structures over children less endowed.

2. Fluency of the use of language does not seem to be a reliable index of maturity of the use of language.

3. Children use great structural variation in their oral language which they are unable to carry over into their written language.

4. Even as early as fourth grade, in narration children use the chief structures of the English language proportionately in a way that approaches the use of the adults who write nonfiction for books, magazines, etc.

5. The written language of the children of this study is inferior structurally to that of most of their textbooks, but not much. It is much more structurally varied than that of the textbooks designed for slow readers.

6. Sixth-grade textbooks do not use language in a way which is much superior structurally to the language of fourth-grade textbooks.

7. None of the textbooks give attention to consistent development of sentence structure.

8. There seems to be some relationship between the use of clauses and phrases and the paragraph comprehension of children in Grade 4; but this relationship is not clear at the sixth-grade level.

In general, Riling found that the oral language patterns used by the children in her sample were quite similar to those reported by Strickland (1962). Differences could be due to any number of these factors: (1) the composition of the samples was not the same; (2) the stimulant to conversation was not the same—Strickland sought open discussion from two or three children while Riling steered the oral response from children by having them tell a story about a picture; (3) the setting was not the same —other children were the focus in Strickland's tape sessions while just an adult was present at Riling's; (4) the language samples were collected in different parts of the country (Indiana and Oklahoma).

Despite some expected differences among the studies, Riling's work strengthens some of the tentative findings of both Strickland (1962) and Hocker (1963). First, relatively large samples of children's oral language can be recorded and analyzed according to the two-level scheme devised in 1959 (Strickland, 1962; Hocker, 1963; and Riling, 1965). Second, children's textbooks do not give much attention to a consistent development of sentence structure (Strickland, 1962; Riling, 1965). Third, it seems as though the kinds of structures that children use are affected by certain aspects of the situation in which the oral language is produced (Hocker, 1953; Riling, 1965). As previously discussed, Hocker emphasized the relative degree of informality and formality about the location of the child's speech—schoolroom vs. playground. Riling looked at a slightly different aspect by emphasizing the purpose for the child's speech; she suggested that further research should consider the association between structure and purpose of the language.

In a work that differed considerably from the methodology of the three studies discussed above, Loban (1963) reported on the longitudinal analysis of language used by children in kindergarten and the six years of elementary school. This study was continued through high school, but only the section relevant to elementary school is discussed here. Loban was particularly concerned with (1) defining the relationship among children's use of oral, written, listening, and reading language, (2) developing fundamental methods of analysis in the study of children's language, (3) locating significant features of language that are worthy of further study, and (4) determining predictable stages of growth in language.

For the study, eleven kindergarten classrooms were matched with family backgrounds that were similar to the general population in the city of Oakland, California; stratification was based on sex, racial background, socioeconomic status, and intellectual ability. A total of 338 children were included in the sample.[4]

Each year, samples of speech, writing, and reading were collected for each child under controlled conditions that were similar for all subjects. In addition, more extensive information was gathered for two subgroups which represented extreme deviations from the mean: a group of 30 subjects who were very high in language ability, and a group of 24 subjects who were very low in language ability. The criterion for determining the subgroups was two or more standard deviations from the mean of two measures which were given equal weight, vocabulary and teacher ratings. Although Loban was concerned with children's written language, the following discussion is limited to those procedures and findings that pertain directly to knowledge of children's oral language and its relation to reading.

The procedure for collecting the oral language samples was more like Riling's (1965) than Strickland's (1962) or Hocker's (1963) in that children were interviewed individually. The examiner played a more definite role in stimulating conversation than in Riling's study, however, as the interview was divided into two parts: (1) the children were asked about "playmates, games, television, illness, and wishes [p. 3]"; then (2) the children were shown a set of six pictures (the same ones being used for all subjects) and asked to discuss "what they saw in each picture and what they thought about each picture [p. 4]." The responses were recorded on an Audograph and then transcribed into typewritten form for analysis.

---

[4] In similar studies, Strickland (1962) used 575 children in Grades 1-6 and Riling (1965) used 300 in Grades 4 and 6.

The oral language samples were analyzed according to the two-level scheme previously discussed, in which the speech sample was segmented by intonation patterns and then divided into syntactic units which were examined for types of subordination. Since Loban has stated that all communication units (what Strickland called "patterns of language structure") can be classified as one of nine patterns, he limited his tabulation of pattern-types to the following list:

Pattern	Symbol	Examples
one	1 2 or 1 ②	Mary eats. (or) Mary is home.
two	1 2 4	Mary eats strawberries.
three	1 ② 5	Strawberries are berries. Strawberries are good.
four	1 2 3 4	Mary threw the dog some biscuits.
five	1 2 4 6	They elected Mary president. They thought Susie conceited.
six	(1) (2) 1	Here is Mary. There are four houses on Lime Street.
seven	Questions	How does he do it? Is he here?
eight	Passive forms	Strawberries were eaten by Mary.
nine	Requests, commands	Go home. (or) Let us go home.
(ten)	Partials	Any incomplete unit. (This is not actually a pattern like the preceding nine patterns.) [pp. 14-15]

From inspection of the list, it is clear that the distinctions between 1 2 M_2 and 1 2 made in Strickland's work (1962) were not made in his analyses; the former pattern was considered an elaboration of the latter.

In addition to this gross analysis of the language samples, six analyses were performed:[5]

1. Classification of speech in terms of its function.

Eight categories were used for this purpose: facts and unelaborated perceptions, interpretations, personal associations, tentative statements or suppositions, generalizations, irrelevancies, direct questions, and figurative language. As Loban pointed out, his scheme was useful for this particular study, but under other data-gathering situations, it might not be comprehensive enough.

2. Classification of oral language style.

In the study, eleven features of oral language style were identified and considered along a continuum which was scaled from 1 to 3:

fluent	to halting
deliberate	to impulsive
coherent	to incoherent, disorganized
energetic	to listless, weak, tired
laconic	to loquacious
expressive	to flat, expressionless
mature	to babyish
distinct, clearly articulated	to blurred, indistinct, mumbled
conventional in usage	to unconventional in usage
ready in response	to slow in response
relaxed	to tense, strained

[p. 17]

Some inconsistencies in the scale can be noted. Most of the language features in the left column are clearly positive attributes which move along the continuum to clearly negative characteristics in the right column. Two of the terms in the left column, however, are not categorically either positive or negative. Is "laconic" a more desirable trait than "loquacious"? Does being "conventional in usage" (rated a 1, the highest score) mean being better than "unconventional usage" (rated a 3, the lowest score)? Perhaps the latter includes the use of metaphor or imagery, which is not generally considered conventional. Loban may have clarified these points to those who were responsible for rating the children's language.

---

[5] Analyses 1, 5, and 6 were made only on data from the two subgroups. Analysis 3 was made only on data from the two subgroups and a small random sample of the total group.

3. Amount of subordination.

4. Difficulties with conventions in usage and grammar.

These categories for nonstandard usage, syntax, and grammar were items involving (1) the use of verbs, pronouns, prepositions, and conjunctions, modifiers, or nouns, and (2) omission or repetition.

5. Vocabulary measured by word frequencies.

6. Vocabulary measured by diversity.

In addition, several ratings were obtained. Each year, the teachers rated the subjects on these language measures: (1) amount of language, (2) quality of vocabulary, (3) skill in communication, (4) organization, purpose, and control of language, (5) wealth of ideas, and (6) quality of listening.[6] An index of reading ability was obtained by having the investigator assign weights to each book that the child completed in a year during the primary grades. From the fourth grade on, results from the Stanford Achievement Test provided information about the subject's reading performance.

---

[6] These measures are clearly subjective in nature. Where, for example, was "amount of language" rated—in the classroom or on the playground? For some children, there may be a drastic reduction in language produced in what may be considered the threatening environment of a classroom.

Some of the pertinent findings of this massive study are summarized below.

1. Over a period of seven years, the subjects increased the amount of language used in the same controlled situation and increased the smoothness of their expression. Members of the low subgroup evidenced more difficulty in using and controlling language patterns; they had less to say and had more difficulty in saying it.

2. The subjects varied little in their use of the basic language patterns; the distinction between the high group and the other subjects was defined as greater dexterity in varying elements within the patterns. By Loban's definition, this means such things as the substitution of word groups for single words, and the use and placement of movable elements of the pattern.

3. Positive relationships between oral language and written language, oral language and reading, and reading and written language were found. Loban stated that subjects who were good readers at the end of the third grade were those children who had "ranked high in oral language" (p. 69) in kindergarten and the first three elementary school grades.[7]

In sum, Loban felt that a major purpose of the study was attained—a scientific method for analyzing children's language and for locating certain characteristics that are important for further investigation. The variety of techniques used in assessing the nature of children's language, the use of structural linguistics in addition to a limited use of transformational grammar on two subjects, and the longitudinal nature of the work make Loban's study unique in its contribution to knowledge about children's language.

---

[7] The meaning of ranking "high in oral language" is not clarified in the context in which this statement appears. One assumes that Loban was referring to an average of the six teacher ratings previously discussed in this chapter.

Some weaknesses in this undertaking were inevitable, however. Perhaps in dealing with language assessment it is difficult to use measures that are completely free from subjective bias; but the questionable method of determining reading proficiency by the number of books read could easily have been made more reliable and valid. Standardized reading achievement tests are available for the primary grades; these would have been a more reasonable measure of relationships between certain language characteristics and reading at the primary level.

An area that did not receive any consideration was the school environment. It would be helpful to know, for example, whether or not the teachers were equally qualified in those classrooms that contained students of predominantly good or poor ability. One might ask if children in the low group performed differently from those in the high group because they had poorer teachers, or because they went to less well-equipped schools. Loban did not discuss these important variables.

Thus far in the review, it should be apparent that each of the four researchers (Strickland, 1962; Hocker, 1963; Riling, 1965; and Loban, 1963) used a method of oral language analysis derived primarily from structural linguistics which provides, among other things, the description of "arrangements of the patterns of grammatical structure [Fries, 1963, p. 73]." In a recent study, O'Donnell, Griffin, and Norris (1967) were influenced more by methods of analysis derived from transformational grammar. In particular, they were concerned with sentence-combining transformations, described in the study as the process of "converting a pair of sentences into a single sentence by embedding one in the other [p. 35]." The general purpose of their investigation was to analyze samples of speech and writing from elementary school children so that any differences in grammatical complexities could be identified and then compared for boys and girls at different grade levels.

Thirty children in each of Grades K–3, 5, and 7 comprised the sample; there were an unequal number of boys and girls at each grade level. All children were pupils in one school in Murfreesboro, Tennessee, with the exception of the kindergartners, who were enrolled in private kindergartens. The language samples consisted of responses to two short movies. The procedure was the same for all subjects: First, children viewed the movies, shown without sound, in groups of threes and were then asked to tell the story to an examiner in private. A series of standard questions was asked after the child gave his interpretation of the movie. All oral responses were recorded by tape. Second, children in Grades 3, 5, and 7 were then asked to write the story of the movie and to write answers to the questions. The purpose of the questions was to elicit a different kind of response than just narration. It will be remembered from the preceding discussion that other researchers handled this problem differently; Hocker (1963) obtained language samples in purposely uncontrolled situations, while Riling (1965) made every effort to restrict the responses to narration only.

A transcript of each child's responses was made, and the typescripts were subsequently segmented into minimal terminable units (T-units), which have been demonstrated to be more reliable units of analysis than the phonological units previously used in studies of this nature.[8] Hunt (1965) has discussed this unit as being "the shortest segments which it would be grammatically allowable to write with a capital letter at one end and a period at the other, leaving no fragment as residue." Thus, "The lady with whom I played tennis last night was never intended to be and is still not intended to be my wife" is one T-unit, where, "She looked and he looked" is two T-units.

---

[8] See note 1.

Once the samples had been segmented, each T-unit was typed on an analysis sheet and analyzed according to the following procedure: (1) the "language pattern" of the main clause was determined (e.g., noun-verb-object, noun-verb, etc.); (2) sentence-combining transformations were classified under three headings—transformations producing nominal constructions, those producing adverbial constructions, and those producing coordinate constructions within T-units.

The data were presented in terms of group means or rates of occurrence per 100 t-units. Statistical treatment included analysis of variance to test the mean differences in the frequency of certain grammatical structures for boys and girls at each grade level and for oral and written language.

Among the pertinent conclusions of the study were the following:

1. Use of the mean word-length of T-units provides a valid and simple measure of the development of syntactic control of children.

2. The number of sentence-combining transformations per T-unit is significantly greater for each advance in grade level, and it is greater in written language in Grades 5 and 7.

3. In the higher grades (5 and 7) control of written syntax is far more evident than control of syntax in oral language.

4. Oral expression progresses fastest between kindergarten and the end of first grade, and again between the end of fifth grade and the end of seventh grade.

5. There is a positive correlation between advance in grade level and increased word-length of response for elementary school children.

6. With respect to differences between oral and written language, the length of T-units in oral language was significantly greater than in written language for Grade 3; but in Grades 5 and 7 the written T-units were longer, though not significantly so.

7. No distinct difference between boys and girls was evident in speech. In writing, however, girls in Grades 3 and 5 were superior.

The O'Donnell et al. study raises several questions about the "unimpressive progress in syntactic control in the middle grades [p. 100]." The investigators suggested that research on the production and interpretation of grammatical structures could lead to improved school programs; children's ability to handle structures in their language could be developed for greater flexibility of expression. Besides suggesting several areas for further research, perhaps one of the major contributions of this study is its application, for the first time, of a refined measure—the T-unit—to analyses of children's oral language. The results of future studies that are based on this measure should be more dependable than those using the less easily determined phonological unit.

Five broad studies of children's oral language have been reviewed for the period 1959-1968 which followed development of a relatively new scheme for analyzing oral language samples. Some of the major contributions of the studies have been noted throughout the preceding review. They can be summarized again briefly as follows:

1. The studies demonstrated that children's language can be analyzed by methods that offer much potential for describing a variety of language characteristics. The two kinds of schemes, one from structural linguistics and the other from transformational grammar, make it possible to work from the kindergartner's language to the more complex language of the sixth graders and yet compare the two along the same measures. The studies differed, however, in the amount of attention paid to careful analysis of

subordination, or the Level II analysis. The emphasis in Strickland's work (1962) was primarily the categorization of language into language patterns, where Loban (1963) provided a more detailed analysis of the kinds of subordinations used, and O'Donnell, Griffin, and Norris (1967) studied the kinds of sentence-combining transformation used.

2. Of the three researchers who investigated the relationship between reading comprehension and select characteristics of oral language, each provided evidence that some relationship does exist. Children rating high in reading ability seemed to be more flexible in their use of basic language structures (Strickland, 1962; Riling, 1965; Loban, 1963).

3. A previously unnoticed indication of oral language maturity was made possible by the use of transformational grammar, by which means it was discovered that children's use of sentence-combining transformations increases across grades (O'Donnell, Griffin, and Norris, 1967).

4. In three of the five studies, the investigators concluded that the kinds of structures children read in textbooks are unlike the structures in their own speech (Strickland, 1962; Hocker, 1963; and Riling, 1965). The implication is for research specifically directed to the question of whether or not children would benefit from material that was written more like the way they speak. The present study was designed to investigate this particular problem.

Experimental Studies Using Written Material Based on an Analysis of Children's Oral Language

Analyses of the language used in materials written for children have been done in an attempt to relate children's use of oral language to the language they read (e.g., Davis and Seifert, 1967; Strickland, 1962; Riling, 1965). Research on readability has defined various structural elements, like the number of simple sentences or sentence length, that are more or less difficult for children to read.[9] The study discussed in this section, however, is the only one, to the investigator's knowledge, that has attempted to experiment with children's materials written according to characteristics of children's oral language patterns.

Ruddell (1963) designed a study to test two basic hypotheses:

> The degree of comprehension with which written passages are read is a function of the similarity of oral patterns of language structure to written patterns of language structure in reading passages....
>
> Comprehension scores on reading passages utilizing high frequency patterns of oral language structure will be significantly greater than the comprehension scores on reading passages utilizing low frequency patterns of oral language structure [p. 35].

He selected oral language patterns from Strickland's study (1962) which had occurred with either a high or a low degree of frequency at the fourth-grade level, and from them he constructed six 254-word passages, three that used frequent oral language patterns, and three that used infrequent oral language patterns. The patterns appeared in direct proportion to the number of times they were used in the oral language of fourth graders (Strickland, 1962). The variables of content, length, style, and readability were controlled, although the content of the three passages written with frequent oral language patterns was only similar, not identical, to that of the other three passages.

---

[9] Summaries of readability research that has dealt with language variables other than the one in this study include those by Gray (1947), Betts (1949), and Ruddell (1963).

Subjects for the study were randomly selected from the fourth grade population of Bloomington, Indiana, where Strickland had recently conducted her study of the oral language patterns of elementary school children (1962). Variables of sex, IQ, mental age, chronological age, and socioeconomic background of the subjects were considered.

Reading comprehension of the passages written with frequent and infrequent oral language patterns was ascertained by the cloze procedure, by which every fifth word in the passages was deleted and a blank space left in its place. The reader's task was to fill in the blanks with the deleted words. The tests were scored by two different procedures. In one, the exact same missing word had to be supplied in order for the response to be considered correct; in the second method, a synonym was scored correct.

The subjects were randomly assigned to read either the three frequent oral language passages or the three infrequent oral language passages first. By the end of two testing sessions, each subject had read all six passages.

Reliability coefficients obtained by the split-half method with odd and even numbers, and then corrected by the Spearman-Brown Formula, ranged from .851-.919 for the individual tests. Ruddell reported that all reliability coefficients were significant at the .01 level. A single measure of validity was obtained by correlating subjects' cloze comprehension scores with scores from the Paragraph Meaning section of the Stanford Achievement Test (Intermediate I, Form X_R). The correlations, which ranged from .609 to .738 on individual tests, were reported to be significant at the .01 level.

To test the first hypothesis, a one-way analysis of variance with a repeated measures design was used. Scores from the six passages were treated in the analysis. For both scoring methods, the F ratios were significant at the .01 level. Ruddell concluded that reading comprehension is a function of the similarity of oral and written patterns of language structure.

The second hypothesis was also tested by a one-way analysis of variance with a repeated measures design. For this analysis, scores from the passages written with frequent patterns were pooled and scores from passages written with infrequent patterns were pooled. Again, the F ratios were significant at the .01 level for both scoring methods used. Ruddell concluded that reading comprehension of material that utilizes frequent oral language patterns will be significantly greater than reading comprehension of material that utilizes infrequent oral language patterns.

The relation of six variables to reading comprehension of the frequent and infrequent passages was determined by using a two-way analysis of variance with a repeated measures design. Significant differences at the .01 level were found in comprehension scores of reading passages that utilized frequent and infrequent oral language patterns and the six variables of (1) occupation of the father, (2) education of the father, (3) education of the mother, (4) IQ, (5) mental age, and (6) chronological age. No significant differences were found between performances of girls and boys. Ruddell indicated, however, that boys had a relatively more difficult time with the passages written with infrequent oral language patterns than did the girls.

Although Ruddell was careful to point out that the cloze procedure has been found to be a valid measure of reading comprehension, he did not evaluate its potential bias in a study which dealt with reading comprehension of select language patterns. It seems reasonable to hypothesize that filling in blanks in familiar language structures such as the noun—verb—object pattern is not as difficult a task as supplying words for language structures that vary the position and type of certain elements in it. For example, a characteristic of the infrequent oral language patterns in his passages was their use of movables, which have been demonstrated to be associated with more mature oral language in children (Loban, 1963). Thus although children might have been able to comprehend a sentence that comprised one or more movables, they might have had difficulty supplying missing words merely because the type of movable and its location were unexpected in that particular sentence.

It would have been informative if Ruddell had done an analysis of the location of each child's mistakes in the frequent and infrequent language patterns to see whether there were some relationship between the type of pattern and the mistake made. One might ask, for example, whether the preponderance of errors in the infrequent patterns occurred within the movable elements rather than in the slots of the sentence.

Summary

In the single study that focused on readability in terms of children's oral language, Ruddell (1963) wrote passages composed of either highly frequent or infrequent oral language patterns of fourth graders, as determined by Strickland's study (1962). He constructed cloze comprehension tests for each of his six passages and administered them to 100 fourth graders. A major finding was that reading comprehension of the passages that comprised frequent oral language patterns was significantly greater than comprehension of the passages that comprised infrequent oral language patterns.

The study was limited to fourth grade, and its use of the cloze procedure to measure reading comprehension could conceivably have affected the results. The findings of this significant work should therefore be confirmed by researchers who investigate the effect of select language patterns by means of a different reading comprehension task, and at different grade levels. The present study was designed to do this.

II

METHOD

Five areas of the study are discussed in this chapter: (1) overview of the study, (2) development of the reading comprehension tests, (3) subjects, (4) data collection, and (5) data analysis. Estimates of reliability and validity for the instruments devised by the investigator were obtained simultaneously with the results. Consequently, a discussion of these two characteristics of the tests is deferred to Chapter III.

Overview of the Study

The general design of this study was a 2 X 2 X 2 factorial with two treatments (Test A and Test B), two grade levels (Grade 2 and Grade 4), and two sexes. Subjects for the study were whole classrooms of second and fourth graders in the two schools made available to the investigator. For each subject, the following data were collected: (1) reading comprehension scores on Test A, which comprised frequent oral language patterns, and Test B, which comprised infrequent oral language patterns; (2) a grade placement score on the Paragraph Meaning section of the Stanford Achievement Test (hereafter referred to as the Stanford subtest), Form Y, Primary II for the second graders and Intermediate I for the fourth graders; (3) a total IQ score from the Lorge-Thorndike IQ test, administered by the schools during the second grade; (4) the occupation of the head of the household; and (5) sex. Variables 3 and 4 were not considered in the analyses; they were used only for descriptive purposes to characterize the sample of the study more fully.

Development of the Reading Comprehension Tests

To determine whether or not children comprehend material written with frequent oral language patterns better than they comprehend material written with infrequent oral language patterns, two distinct reading comprehension tests were constructed: Test A comprised language patterns that appear frequently in the oral language of both second and fourth graders (Strickland, 1962); Test B comprised language patterns that appear infrequently in the oral language of both second and fourth graders (Strickland, 1962).

In prior research, Ruddell (1963) wrote 254-word passages which used select oral language patterns in direct proportion to the number of times they appeared in children's speech.[1] The use of a prescribed number of language patterns in each passage resulted in material that was sometimes stilted and contributed.[2] To avoid this problem in the present study, reading comprehension was ascertained for individual sentences that were unrelated to each other; the proportion of each language pattern used in the tests, however, was still controlled. The decision to use unrelated sentences rested on three assumptions:

---

[1] This study was reviewed in Chapter I.

[2] To clarify this point, these sample sentences are quoted respectively from Ruddell's Passage IA and IIB: (1) "Man must learn many new things about space and he plans future space flights" (p. 139), and (2) "After catching fish a Japanese fisherman sells from a big water tub and while in the water the fish will live" (p. 146).

1. In a lengthy passage it is difficult to check the degree of comprehension of each sentence. When working with language patterns that differ in frequency, the variable under careful scrutiny in this study, it is desirable to measure reading comprehension of each sentence.

2. Reading a series of unrelated sentences rather than a lengthy passage, such as those used by Ruddell (1963), was assumed to be better suited to the abilities of the younger subjects in the study.

3. Although second and fourth graders typically read short stories or connected paragraphs rather than isolated sentences, comprehension of a unit as small as a sentence is frequently required in the classroom.

On the basis of these considerations, the tests were a series of sentences. Briefly, three pictures were drawn for each sentence. The reader's task was to select the one picture that "told about" the sentence the best and then draw a line from the sentence to the picture. The procedure for selecting the patterns and exerting control over extraneous variables is outlined below; a more complete discussion of how reading comprehension was measured by multiple choice items appears later in the chapter.

Selection of the Patterns

There have been several recent descriptions of the structures that appear in children's oral language,[3] but for two reasons only Strickland's findings (1962) were used. First, Loban (1963) defined nine patterns into which all oral communication units can be classified.[4] He determined the degree to which each one appeared at various stages of language development of elementary school children. However, because widely different variants of each pattern are present in children's language (Loban, 1963; Strickland, 1962; Riling, 1965), the present study was not confined to Loban's list. Second, in similar studies, Riling (1965) and Hocker (1963) used Strickland's work (1962) as the impetus for their research. They were not as comprehensive; the sample for the former comprised fourth and sixth graders while for the latter it comprised only first graders. Strickland's study (1962), then, with its broad description of structures in children's oral language was the sole source of information about the nature and frequency of children's oral language patterns.

---

[3] Five recent studies that focused on such descriptions were reviewed in Chapter I.

[4] Some examples are the subject—verb—direct object pattern (1 2 4) and the subject—verb—indirect object—object pattern (1 2 3 4). See the review of Loban's study in Chapter I for a complete list of these patterns.

One of the major interests in the present study was the comparison of reading comprehension scores of second and fourth graders. Two alternatives were available in constructing the materials: either write a different set of materials for each grade, or write a set of materials for second graders and give them to subjects at both grade levels. The first alternative was discarded because obtained grade level differences in comprehension of select oral language patterns would have been confounded by all the factors that differentiated the tests at two grade levels. The second alternative was accepted; control over variables of content, length, grammatical complexity, and vocabulary—control that was essential if the focus was to be language pattern difficulty—could be identical for all subjects.[5]

---

[5] In a recent study in which the quality of main idea statements made by elementary school children in Grades 1–6 were compared (Barrett and Otto, 1968), materials were constructed in two ways: (1) a short paragraph was written at the first grade level and given to half the subjects across grades and (2) the paragraph was gradually made more difficult so that the readability level increased for each succeeding grade level. The authors reported significant differences in the main idea statements of succeeding grade levels regardless of whether or not the first grade or the appropriate grade-level paragraph was read by the subjects. From this finding one could reasonably infer that the main idea task was not adversely affected by having older subjects read the presumably easier first grade paragraph. This finding is considered relevant to the decision to give the same tests across grades in this study.

Because the same materials were to be used, a major problem was selecting language patterns from Strickland's study (1962) that appeared in approximately the same rank order of frequency in both second and fourth grades. This was necessary to assure that the meaning of "frequent" and "infrequent" was the same for subjects at both grade levels. This consideration superseded all others until the requirement was met.

Patterns for Test A were selected from among those that ranked the highest in frequency at the second and fourth grades in the Strickland study (1962). Patterns that were ranked lower than tenth in frequency appeared less than one-sixth the number of times the most frequent pattern at either grade level appeared; thus it seemed reasonable to designate as infrequent any pattern lower than tenth in the rank order. The list from which the patterns for each test were drawn is presented in Appendix A.

Three additional criteria were considered in determining which patterns were to be used:

1. Any pattern that ended with a T marker, a conjunction indicating that another pattern should follow (e.g., 1 2 4 + T), had to be followed by a pattern beginning with a T marker, a conjunction indicating that another pattern should precede it (e.g., T 1 2 4).

2. Since two of the most frequent patterns did not contain a T marker, or more than one main clause (1 2 4 and 1 2b 5), an attempt was made to include a similar number of infrequent patterns with only one main clause so that neither test would be overburdened with lengthy sentences. Because it is generally agreed that sentence length influences readability, control of this variable seemed necessary.

3. Patterns were to appear in approximate proportion to the number of times they appeared in children's oral language.

The patterns selected for Test A are presented in Table 1 and patterns selected for Test B are presented in Table 2 along with information about the rank order and frequency for each grade level. As shown in the tables,

Table 1

The Rank Order and Frequency of Oral Language Patterns Selected for Test A

Oral Language Patterns	Grade 2		Grade 4	
	Rank Order	Frequency[a]	Rank Order	Frequency[b]
1 2 4	1	309	1	202
T 1 2 4	2	264	2	156
1 2 4 + T	3	189	3	123
1 2b 5	4	130	4	108
T 1 2 4 + T	6	74	5	84

[a] These figures are taken from data on the 100 second graders in Strickland's study (1962).

[b] These figures are taken from the data on the 75 fourth graders in Strickland's study (1962).

several differences between the patterns of Test A and Test B are apparent. First, the rank orders of the infrequent patterns were less similar across grades than the frequent patterns. Since the difference in rank order was never greater than two, however, the discrepancy was considered insignificant for purposes of this study. A second characteristic of the infrequent patterns was their diversity, which contrasted sharply to the 1 2 4 component in all but one of the frequent patterns. This diversity was expected; research has indicated that flexible use of the basic language structures is a trait of mature and therefore more individual oral language in children (Loban, 1963).

A third difference between the two tests was the larger number of patterns selected for Test B; five were used in Test A compared to six in

Table 2

The Rank Order and Frequency of Oral Language Patterns Selected for Test B

Oral Language Patterns	Grade 2		Grade 4	
	Rank Order	Frequency[a]	Rank Order	Frequency[b]
T 1 2 M_2	17	32	17	25
1 2 M_2	17	32	19	22
T 1 2 4 M_2	19	29	20	21
M_3 1 2 M_1 +T	20	28	21	19
1 2 3 4	23	23	24	15
1 2 4 M_1 +T	25	21	23	17

[a] These figures are taken from data on the 100 second graders in Strickland's study (1962).

[b] These figures are taken from data on the 75 fourth graders in Strickland's study (1962).

Test B. Adding a sixth to Test A would have meant including a pattern of much lower frequency of occurrence than the five top-ranking patterns that were used. A pattern with a T marker was added to Test B—making two patterns that ended with a conjunction and two that began with one—so that an approximate similarity of the oral and written language could be maintained.

Determination of the Proportion of Patterns to be Used

Children's oral language patterns differ greatly in the frequency of their occurrence. To assure some similarity between the structure of children's oral language and the written language of the tests, each pattern was used in approximate proportion to the number of times it appears in children's oral language (Strickland, 1962).

Several steps were necessary to determine how many times each pattern should be used. As demonstrated previously in Tables 1 and 2, the frequencies and rank orders of the patterns differed from grade to grade. The general procedure was to compute the proportions separately for each grade and then average them. More specifically, these computations were made for each test:

1. The frequencies of all the patterns within the test were added. This process was repeated separately for Grades 2 and 4.

2. The frequency of each pattern was then divided into the resulting sum, yielding an index for each pattern.

3. To find the average index for each pattern across Grades 2 and 4, the second and fourth grade indices were added and divided by two. This final number was used to determine how many times each pattern would appear.

An example is provided for clarification of this process: In Test A, the frequencies of the five patterns totaled 966 at the second grade level. This total was divided into the frequency of the 1 2 4 pattern, 309, to determine its index. The resulting number, .32, was the index, or proportion, for the 1 2 4 pattern.

Derived in the same manner, the index for the 1 2 4 pattern at the fourth grade level was .30. Averaging .32 and .30 resulted in the final index, .31. Thus 31% of the frequent patterns in Test A should have been 1 2 4 (noun–verb–object).

The indices computed for each pattern in Test A and Test B are summarized respectively in Tables 3 and 4 along with the actual proportions used.

Table 3

Index of Proportions of Frequent Oral Language Patterns in Material Written for Second and Fourth Graders

Frequent Oral Language Patterns	Index (The average between proportions determined separately for Grades 2 and 4)	Actual Proportions Used
1 2 4	.31	.32
T 1 2 4	.25	.23
1 2 4 +T	.19	.23
1 2b 5	.15	.11
T 1 2 4 +T	.10	.12

Table 4

Index of Proportions of Infrequent Oral Language Patterns in Material Written for Second and Fourth Graders

Infrequent Oral Language Patterns	Index (The average between proportions determined separately for Grades 2 and 4)	Actual Proportions Used
T 1 2 M_2	.20	.15
1 2 M_2	.19	.18
T 1 2 4 M_2	.18	.20
M_3 1 2 M_1 +T	.17	.18
1 2 3 4	.14	.15
1 2 4 M_1 +T	.14	.16

Control of Select Variables

Since the primary focus of the study was to determine whether or not reading comprehension of the patterns in one test differed from reading comprehension of the patterns in the other test, control over vocabulary, content, and grammatical complexity was essential in both tests. Furthermore, since the same tests were given to two grades, it was necessary to demonstrate that the three variables were controlled across grades so that the research task did not inadvertently bias the performance of one grade over the other. The manner in which the influence of each variable was defined and controlled is discussed below.

Vocabulary control. Because the same tests were read by children in second and fourth grades, the intent was to make the actual decoding process as uncomplicated as possible for the former so that performance did not depend on familiarity with the words. Two types of controls were imposed. First, the primary source of vocabulary for both tests was the Stone List (Stone, 1953), a revision of the Dale List of 769 Easy Words. Words that did not appear on the list were given to second and fourth graders in a pilot study;[6] if a clearly larger number of the former misread a word, then a Stone List word was substituted in the final version.

A second kind of control was inherent in the fact that the subjects at both grades were directed to raise their hand if they needed help with any words they did not know. (See Appendix B for a copy of the Directions.) From pilot studies that preceded final testing, it was clear that second graders in particular relied on and welcomed help with words if they needed it. Even the simplest words were requested by some of the children—words that appeared on the Stone List.

---

[6] The words were listed in random order. In a separate testing room, children from two second grade classrooms read the words to the examiner. Mistakes were recorded on a separate list.

Though it was not practical to use the same words in each test,[7] an attempt was made to overlap the vocabulary as much as possible. Thus two additional constraints were imposed: (1) the same number of "hard" words (words not appearing on the Stone List) were used in both tests; and (2) a type-token ratio which has been used in prior research (Loban, 1963; Kean, 1967) was made as similar in each test as it was possible and practical to do. In this ratio, the number of different words used (type) is related to the total number of words used (token).

In Table 5 the "hard" words for each test are listed, and in Table 6 the type-token ratios for Test A and Test B are presented.

Content control. Because the vocabulary was drawn largely from the Stone List (Stone, 1953), the content of each test was limited to objects and actions that were within the reading vocabulary of a typical second grader. To attempt to write reasonably interesting test materials for both grades, the following criteria were used to determine the content of sentences within the tests:

1. Highly implausible situations which required knowledge of difficult relationships of facts were avoided in order to make the concepts and ideas familiar to the second and fourth graders alike.

---

[7] The to be verb, for example, appeared in one pattern in Test A (1 2b 5) and in no patterns in Test B. The verbs were, are, and is followed by the predicate nominative permeated portions of Test A whereas they were not included in Test B. Thus it was not possible to match the vocabulary of something as apparently straightforward and simple as the verbs.

Table 5

"Hard" Words Used in Test A and Test B

Test	Hard Words			
Test A				
(frequent	broken	bone	bite	dried
oral language	goes	kite	larger	mailbox
patterns)	longest	playful	smaller	
Test B				
(infrequent	airport	cream	closely	drums
oral language	goes	glass	lying	melts
patterns)	quickly	boots	threw	

Table 6

Ratio of the Number of Different Words (Type) to the Total Number of Words (Token) in Test A and Test B

Test	Type-Token Ratio
Test A (frequent oral language patterns)	182 (type) / 387 (token) = .47
Test B (infrequent oral language patterns)	204 (type) / 469 (token) = .43

2. Closely related to the first criterion was the use of situations that involved animals and family activities in most of the sentences. In

materials like basal readers, these topics are widely accepted as being of interest to children at the grade levels used in this study. It was assumed that they provided areas of common experience so that the younger subjects were not put at a disadvantage.

The content of both tests was made as similar as possible, but the obvious possibility of using identical situations was avoided for two reasons. First, the tests were given a day apart from each other; memory was expected to interfere with responses made to the second test. Second, the movable elements that characterized the infrequent patterns in Test B concerned details of time, manner, and place—concepts that were not a part of the frequent patterns of Test A. Sentences in Test B, then, frequently involved details illustrated by the movables "above the tree," "now," and "near the door" that could not be incorporated logically into sentences in Test A.

To provide some measure of consistency between the tests, however, the content of each test was categorized. The number of sentences falling into eight groups was tabulated and compared with the number in the other test; unduly large discrepancies between the tests were corrected. The number of sentences falling into the eight categories according to sentence content is summarized in Table 7 for each test. The number of sentences in each category according to picture content is summarized in Table 8. On the basis of the tables, these observations can be made: (1) the two tests dealt with similar content as determined by the written sentences in each test, and (2) the two tests dealt with similar content as determined by the picture content in each test.

Grammatical complexity control. Strickland's study of children's oral language patterns (1962) focused primarily on the type of slots in the patterns (e.g., subject, verb, direct object) and the type and location of

Table 7

Analysis of Sentence Content in Test A and Test B

Category	Number of Sentences Falling in Category	
	Test A	Test B
Animals	13	7
Children (male and female)	4	5
Males	5	8
Females	2	3
Toys and objects	9	6
Nature	2	4
Animals and males	2	2
Animals and females	-	2

Table 8

Analysis of Picture Content in Test A and Test B

Category	Number of Picture Sets in Each Category*	
	Test A	Test B
Animals	13	7
Children (male and female)	4	7
Males	5	10
Females	3	2
Toys and objects	8	3
Nature	2	4
Animals and males	2	2
Animals and females	-	2

*A picture set is defined as the three pictures that comprised the multiple choice items for each sentence.

the movables. How each slot is filled, however, seems to be of crucial importance in distinguishing mature oral language (Loban, 1963). Educational experience and reason readily lead to the conclusion that for children the reading task is complicated by putting phrases and clauses in slots instead of single words. Given the wide range of possibilities for each slot, the decision was made to include different kinds of fillers (e.g., single words, phrases, clauses) rather than limit the selection to single words. On the other hand, it was necessary to avoid overloading the slots with complexities that would bias the performance of the younger subjects.

To balance these two concerns, the preponderance of slots were filled with single words, words modified by one or two adjectives, and compound nouns, verbs, and objects.

Phrases and clauses were used more sparingly with the exception of constructions in movables, many of which were filled with a phrase to avoid needless repetition of single-word adverbs. "In winter" and "at night" are examples of substitutions for single-word adverbs in the infrequent oral language patterns that contained the movable M_3—the adverb of time.

An attempt was made to distribute the kinds of slot fillers among the various patterns within each test. This precaution was intended to avoid overloading the simplest and shortest patterns within each test so that they became proportionally more difficult than some of the longer patterns.

Indices of subordination from prior research (Loban, 1963; Kean, 1967) did not readily lend themselves to distinguishing among the kinds of slot fillers controlled in the present study. For the purpose of convenient classification, the following index was devised. Examples for each of the six categories are provided:

Category	Slot Filler	Example for the Subject Slot	
0	single word	The boy	ran.
1	single word + adjective	The slow boy	ran.
	single word + adjectives	The slow, tired boy	ran.
2	phrase	The boy with the coat	ran.
	phrase + adjective	The boy with the red coat	ran.
	phrase + adjectives	The boy with the torn red coat	ran.
3	compound	The boy and girl	ran.
	compound + adjective	The boy and tall girl	ran.
	compound + adjectives	The short boy and tall girl	ran.
4	clause	The boy who had a coat	ran.
	clause + adjective	The boy who had a red coat	ran.
	clause + adjectives	The boy who had a torn red coat	ran.

It is clear that for other purposes, the index would need to be revised; but it was adequate for its limited use in describing the kinds of grammatical complexities utilized in the present study.

The Index Rating (IR) given to slot fillers in Test A is summarized in Table 9 for each pattern, and the IR for slots in Test B is likewise summarized by pattern in Table 10. From inspection of the subject slot for the 1 2 4 pattern in Table 9, for example, 8 subject slots were filled with single words, 5 with single words + adjective(s), 3 with phrases, 1 with a compound subject, and 1 with a negative.

To make comparisons between the tests more readily apparent, only the slots that patterns in both tests had in common (subject, verb, object) were analyzed; consequently movables and the indirect object of certain

Table 9

Total Index Ratings (IR) Given to Slots in the Oral Language Patterns of Test A

Frequent Oral Language Patterns in Test A	Subject (1)	Verb (2) or (2b)	Object (4)	Predicate Nominative (5)
	IR 0 1 2 3 4	IR 0 3 5	IR 0 1 2 3 4	IR 0 1 2 3 4
1 2 4	8, 5, 3, 1, 1	16, 1, 2	8, 6, 1, 3, 0	
T 1 2 4	8, 6, 0, 0, 1	13, 0, 1	8, 6, 1, 0, 0	
1 2 4 +T	8, 6, 1, 0, 0	11, 2, 1	8, 6, 1, 0, 0	
1 2b 5	3, 2, 2, 0, 0	7, 0, 0		2, 2, 0, 2, 0
T 1 2 4 +T	6, 1, 1, 0, 0	7, 1, 0	4, 4, 0, 0, 0	

patterns in Test B were not included. A summary of the analysis for each test is presented in Table 11, where it is clear that the slot fillers were similar from test to test. It was concluded that the grammatical complexities were adequately controlled across tests.

Measurement of Reading Comprehension

Since comprehension of each sentence in the tests was assessed, more typical kinds of tasks such as selecting the correct sequence of events or a title for a story were not appropriate. Pictures have been used to measure reading comprehension in standardized reading achievement tests (Gates Advance Primary Reading Test, Type APR, Form I) and in recent research on sentence constructions (Ruddell, 1965). In the latter instance, four sentences were listed opposite two pictures. The task was to draw a line from each picture to the one sentence that told about the picture.

Table 10

Total Index Ratings (IR) Given to Slots in the Oral Language Patterns of Test B

Infrequent Oral Language Patterns in Test B	Adverb of Time (M_3)	Subject (1)	Verb (2)	Indirect Object (3)	Object (4)	Adverb of Place (M_1)	Adverb of Manner (M_2)
	IR 0 1 2 3 4	IR 0 1 2 3 4	IR 0 3 5	IR 0 1 2 3 4	IR 0 1 2 3 4	IR 0 1 2 3 4	IR 0 1 2 3 4
1 2 M_2	______	4,4,1,0,1	9,1,0	______	______	______	1,1,4,3,0
1234	______	4,3,2,0,0	7,2,0	7,2,0,0,0	2,3,1,2,1	______	______
$M_3$12M_1	8,0,4,0,1	7,4,1,1,0	13,0,0	______	______	2,0,11,0,0	______
124M_1	______	7,3,0,0,0	7,2,1	______	6,4,0,0,0	3,0,7,0,0	______
T124M_2	______	7,4,1,0,0	11,1,0	______	8,3,1,0,0	______	3,1,5,0,1
T12M_2	______	7,2,0,0,0	7,1,1	______		______	2,1,6,0,0

Table 11

Summary of Total Index Ratings (IR) Given to Slots in Oral Language Patterns Common to Test A and Test B

Test	Subject (1)	Verb (2) or (2b)	Object (4)
	IR 0 1 2 3 4	IR 0 3 5	IR 0 1 2 3 4
Test A (frequent oral language patterns)	33,20,7,1,2 (N = 63)	54,4,3 (N = 63)	28,22,3,3,0 (N = 56)
Test B (infrequent oral language patterns)	36,20,5,1,1 (N = 64)	54,7,2 (N = 64)	16,10,2,2,1 (N = 31)

Note—The numbers in the slots differ according to the nature of the pattern being analyzed. In the 1 2b 5 pattern for Test A, for example, only the first two slots were included for analysis here.

Considering both the relatively large number of sentences to be read in the present study—37 in each test[8]—and the endurance limitations of the younger subjects, it seemed reasonable to require one sentence rather than two to be read for each response. For each sentence, then, three pictures were drawn. One depicted the content of the sentence and was the correct response. Each of the other two pictures held most of the details of the correct picture constant while varying one major detail at a time.

---

[8] The reason for this particular number of items is discussed later in the chapter under the heading, "Final Selection of the Test Items."

An example of the nature of the distractors is provided in Figure 1; the item is from Test B. It was intended that the basic likeness of the three pictures in each multiple choice set would require careful reading of the sentence in order to make a correct response. Reading comprehension in this study, then, was measured by the child's ability to read a sentence and select one of three similar pictures that belonged with the sentence.

The artist was directed to use simple line drawings and to make each of the three pictures similar except for the major change in each distractor. If a boy and girl were to be the subject for a set of pictures, for example, the same boy and girl were used in each picture although their poses might be different depending on the discrimination being required of the reader. Figure 1 illustrates this criterion.

A rigid formula for constructing each distractor was not appropriate, as variety in the content and grammatical complexities necessitated flexible treatment of the sentence. Certain guidelines were established, however, to assure that (1) different kinds of discriminations were included within a test, and (2) the discriminations used in one test were similar to those used in the other test. An attempt was made to emphasize reading comprehension of the particular characteristic of the pattern being read. In the sentence "Sue does give Tom some fish" (the infrequent pattern 1 2 3 4), for example, one of the pictures showed Tom giving Sue some fish while a second picture showed Sue giving Tom one fish. The following types of discriminations were used:

	Variable	Correct Picture	One Distractor
(1)	noun-verb agreement	The dog sees the cat.	The dogs see the cat.
(2)	negative-affirmative	Mr. Brown's new coat comes without a pocket.	Mr. Brown's coat comes with a pocket.

The young boy without a coat hits with a bat.

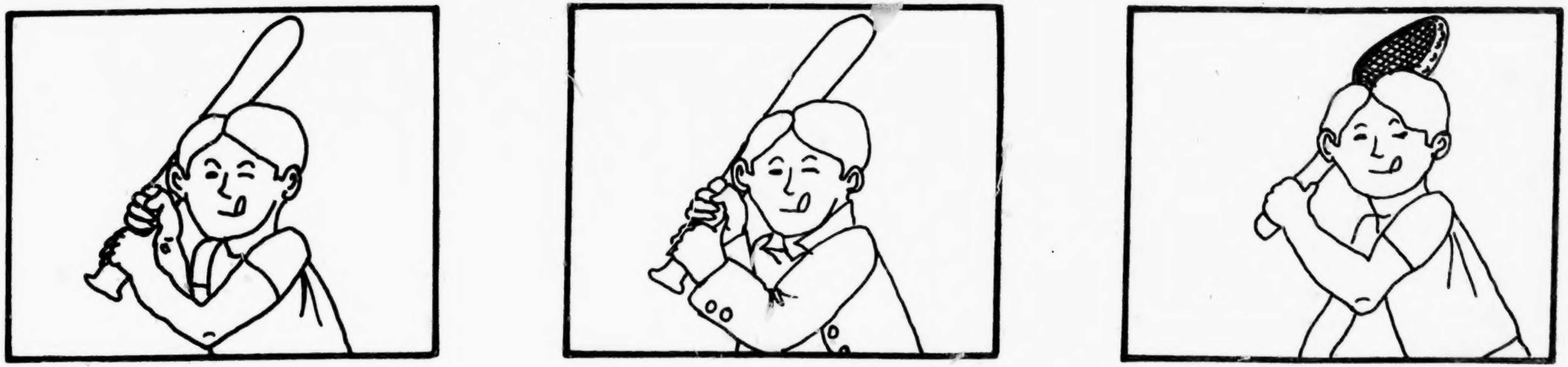


Fig. 1. Sample Item from Test B.

	Variable	Correct Picture	One Distractor
(3)	count nouns	Just Jane's kite has two tails.	Two kites have two tails.
(4)	mass nouns	The fish are eat- the food.	The fish is eating the food.
(5)	verb tense	The boy in the coat will throw the ball.	The boy in the coat has thrown the ball.
(6)	manner*	She draws quickly.	She draws slowly.
(7)	place*	My friend draws a house on the paper.	My friend draws a house on the sidewalk.

*These discriminations, which involved movables, were only relevant for some of the infrequent oral language patterns used in Test B.

Final Selection of Test Items

To obtain some feedback from children and also statistical information about the tests and their items prior to final testing, a pilot study was conducted. Two whole classrooms of second graders and one of fourth graders, the number of classrooms made available in a nearby elementary school, served as subjects. From an inspection of several basal readers at the second grade level, a reasonable length for each of the final tests was judged to be approximately 35-40 items—the average number of sentences in the second grade stories. About one and one-half that many items for each test, distributed proportionally among the patterns within each test, were constructed and piloted. The frequent and infrequent oral language patterns were grouped separately into two booklets, which were preceded by a brief set of directions and a practice sentence. As a check on the clarity and unambiguity of the pictures, subjects were directed to raise their hands if they did not understand a picture. They could also ask for any words that they did not know.

The children were randomly assigned to read either Test A (frequent oral language patterns) or Test B (infrequent oral language patterns) first. A test was read on each of two consecutive mornings of testing for each classroom.

The data for each subject comprised a single score for each of the reading comprehension tests. In addition, to help determine whether or not an item discriminated between good and poor readers, a rating of from 1 to 3 given by the teacher on the basis of reading achievement was obtained. In analyzing the data, a distinction between grade levels was not made due to the small number of fourth graders. By means of the 1604 computer program FORTAP, a Fortran Test Analysis Package by Baker and Martin (1968) at the University of Wisconsin Computing Center, an item analysis was performed. Among the information provided by this analysis were the following: item difficulty (P), item correlation with the rest of the test (R), X_{50}, and Beta. In addition, a Hoyt Reliability Coefficient was obtained for each test, the reliability coefficient being .85 for Test A and .86 for Test B.

A decision was made to weigh the following factors in determining which items were to be considered for the final study: (1) item difficulty, as expressed by the proportion of subjects who responded correctly to the item (P); (2) correlation of the item with the rest of the test (R); (3) the degree to which the item discriminated between good and poor readers, as determined by the teacher ratings for each subject; (4) the degree to which it was felt that replacing one of the distractors would improve the discriminating power of the item considerably. An example of a distractor that functioned poorly in the pilot study is provided in Figure 2; the item, from Test B, was modified so that the middle distractor was replaced by a presumably less obvious one. The revision is presented in Figure 3.

She buys her children some milk and eggs.

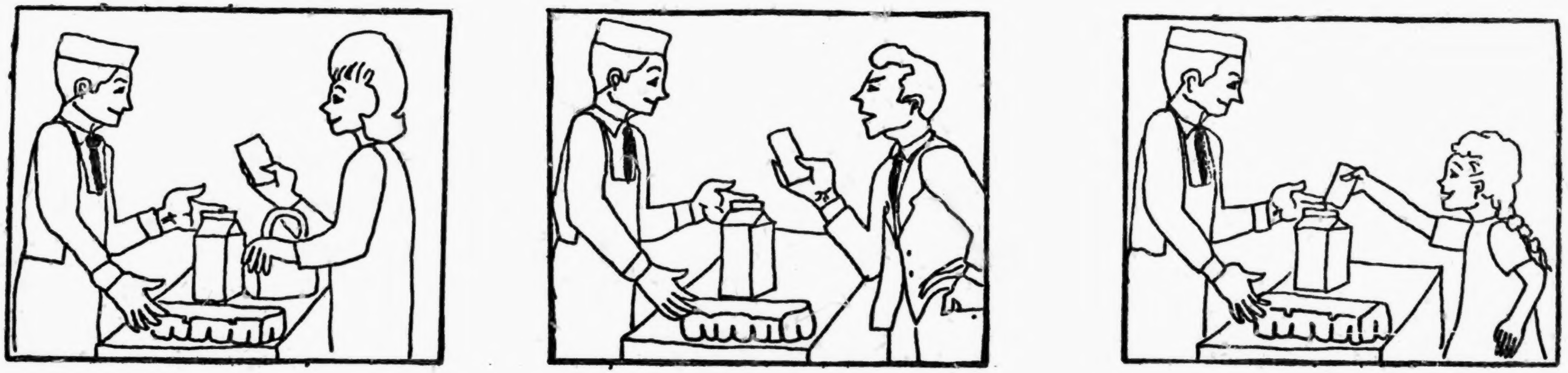


Fig. 2. A Poor Distractor in the Pilot Study from Test B. (The Middle Picture)

She is buying her own children some milk and eggs.

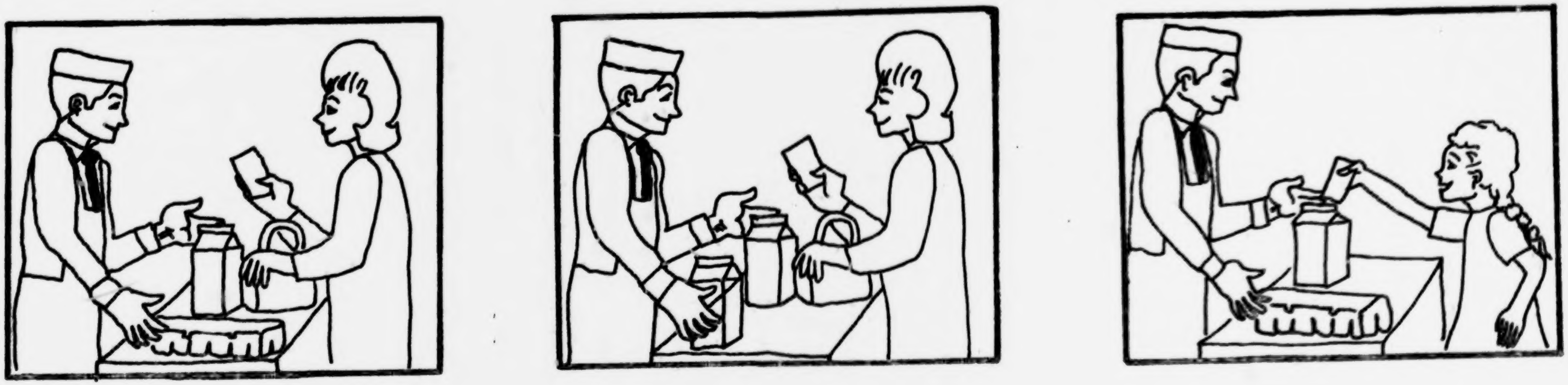


Fig. 3. The Revised Distractor from Test B.
(The Middle Picture)

When the best items were selected according to the above criteria, an attempt was made to improve them so that the tests would discriminate more effectively between the good and poor readers. The mean scores for both tests were considered higher than desirable, being 47.66 for Test A and 43.33 for Test B out of a possible 57. Thus, to make the tests more difficult, one or both of these procedures was used: (1) the distractors were made more difficult by requiring finer discriminations among the three pictures for each sentence, and/or (2) the sentences were made more difficult by expanding them slightly in terms of length, the addition of adjectives, or the addition of more complex grammatical structures like phrases. An example of the former was provided in Figures 2 and 3. An example of the latter procedure is in the sentence in the Test A pilot study, "The monkey has a balloon and a toy car." As it appeared in the final version, the sentence read, "The monkey went and got a balloon and a toy car."

Accordingly, some of the items were revised. A final selection was determined by the proportion that each pattern was to appear in the test.[9] The final number of items that met these requirements was 37 in one test and 42 in the other. Items were deleted from the latter until both tests contained a total of 37 sentences. See Appendix C for a list of the 37 sentences in each test, analyzed in terms of the language patterns used.

Composition of the Final Tests

In sum, each of the two final versions of the reading comprehension tests comprised 37 sentences—two on each page but the last, which contained one sentence. Within each test, the 37 sentences were randomly ordered; and for each sentence the three pictures were randomly ordered. See Appendix B for a copy of Test A and Test B.

---

[9] For a discussion of how the proportion of each pattern was determined, refer to pp. 41-43.

A set of directions preceded each 18-page test. The directions for the first day of testing were somewhat longer than for the second day. During each day, the subjects were told to read a sentence and look carefully at the three pictures under the sentence. They were to decide which picture told about the sentence the best. Then a line from the sentence to the picture was to be drawn. They could raise their hands and ask for any word that they did not know. There was no time limit imposed.

In addition to the directions there were two practice sentences for the first day and one for the second day. These items were devised so that the patterns were neither the frequent nor the infrequent patterns used in the tests. A copy of the Directions for each day of testing is in Appendix B.

Subjects

The subjects were from two elementary schools in Madison, Wisconsin. The schools were chosen because the students were presumed to be reasonably similar with respect to socioeconomic background and spread of ability. School I had five second-grade and four fourth-grade classrooms, and School II had two classrooms at each grade level. All second and fourth graders within these 13 classrooms were designated subjects. The exceptions were children who had repeated a grade. This included 20 second graders and 13 fourth graders. In addition, a teacher asked that one child not be required to participate because of emotional instability. A final total of 163 second graders, 81 girls and 82 boys, and 137 fourth graders, 69 girls and 68 boys, served as subjects.

In order to characterize the sample population, IQ scores and a measure of socioeconomic background were obtained for each subject. Results from the Lorge-Thorndike IQ tests administered by the schools in second grade were used. The mean, range, and standard deviation of IQ scores for all subjects and for subjects grouped by grade, sex, and sex within grade are summarized in Table 12. In the population sample, ability ranged from very low (an IQ of 77) to very high (an IQ of 145); a wide range of ability was also evident in each of the subgroups. The mean IQ of the sample was 105.6, or slightly above average. As the largest difference between the subgroup means and 105.6 was 1.6 (105.6 minus the fourth-grade boys' mean of 104.0), it was concluded that for the sample as a whole and for each subgroup, the subjects could be described as approximately "average" in terms of their mean IQ scores.

Table 12

The Mean, Range, and Standard Deviation of IQ Scores for all Subjects and for Subjects Grouped by Grade, Sex, and Sex within Grade

Group	Mean	Range	Standard Deviation
All Subjects	105.6	76-145	11.7
Second Graders	105.9	82-134	11.2
Fourth Graders	105.3	76-145	12.2
All Boys	104.8	76-145	13.0
All Girls	106.3	77-136	13.6
Second Grade Boys	105.6	84-130	11.8
Second Grade Girls	106.1	82-124	9.5
Fourth Grade Boys	104.0	76-145	12.7
Fourth Grade Girls	106.5	77-136	11.2

The Duncan Socio-economic Index Scale was used to depict the subjects' socioeconomic background. As described by Reiss (1961), a subject's position on the scale is determined by the occupation of the head of the household.[10] Each subject was assigned a score from the index according to the parent's occupation. Then, using Mortenson's method (1966), the occupational levels were combined to yield three socioeconomic levels: a rating of 0–2 was considered low, 3–6 was considered middle, and 7–9 was considered high. The proportion of subjects in each of the three categories is summarized in Table 13. As shown, the spread of occupational levels ranged from the lowest to the highest categories, but over half the sample was rated at the middle socioeconomic level.

Table 13

Proportion of Subjects in Three Levels Defined by the Duncan Socio-economic Index Scale

	Index Rating		
	0-2	3-6	7-9
Proportion of Subjects	.26	.56	.18

In sum, the subjects could be characterized as approximately "average" in terms of their mean IQ score, and the majority could be described as being from the middle socioeconomic level.

---

[10] The validity of this measure was demonstrated by Reiss (1961) when he correlated scores from the Duncan scale with ratings of income and education. The rank correlations between the Duncan scores and income was .85 and between the Duncan scores and education was .83. Reliability of the Duncan scale has not been determined, although a product moment correlation coefficient of .99 was obtained between 1947 scores of the scale from which the Duncan scale was derived and a replication in 1963 (Hodge, Siegel, and Rossi, 1964).

Data Collection

Pupils in each of the seven second grade and six fourth grade classrooms were tested by classroom for three separate testing sessions. During one week of April, 1968, the reading comprehension tests were administered on two consecutive mornings to subjects in each of the thirteen classrooms. During the following week the Stanford subtest was administered. Primary II to the second graders and Intermediate I to the fourth graders. This subtest consists of a series of paragraphs which are graduated in difficulty. From each paragraph, one or two words are deleted. The reader's task is to select the correct word for each blank space from among four choices. The timed test is 30 minutes long for second and fourth grades.[11]

To avoid bias at either grade level, the second- and fourth-grade classrooms were evenly distributed throughout the three testing periods each day. Thus, some second and some fourth graders were tested at either 8:30, 9:30, or 10:30 in the morning. The reading comprehension tests were administered at the same time of day to any one class for each of the two days. Although an hour was allowed per class, the slowest readers took approximately 35 minutes.

All of the testing was done by three experienced elementary school teachers who became familiar with the intent and procedures of the study in a training session prior to testing. Within the two schools, each examiner was randomly assigned to a block of both second- and fourth-grade classrooms; only two examiners were needed to complete the testing for School II, which had four participating classrooms. Subjects who had been absent for one of the testing sessions were tested in small groups during a third week.

---

[11] According to the test manuals for Primary II and Intermediate I (Kelley, Madden, Gardner, and Rudman, 1964), the Kuder-Richardson reliability coefficients and standard errors of measurement for each test are, respectively, .93 and 2.0 at the second grade level and .91 and 4.0 at the fourth grade level; these measures were based on a random sample of 1000 pupils at each grade. The authors discussed content validity in terms of their examination of courses and textbooks to determine the nature of the understandings to be measured.

Data Analysis

To provide answers to the three basic questions of the study,[12] nine hypotheses were tested:

H1: The number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is better than reading comprehension of material written with infrequent oral language patterns is not significantly greater than the number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is not better than reading comprehension of material written with infrequent oral language patterns.

H2: Fourth graders do not do significantly better than second graders in their reading comprehension of material written with frequent oral language patterns.

H3: Fourth graders do not do significantly better than second graders in their reading comprehension of material written with infrequent oral language patterns.

H4: Second- and fourth-grade girls do not do significantly better than second- and fourth-grade boys in their reading comprehension of material written with frequent oral language patterns.

H5: Second- and fourth-grade girls do not do significantly better than second- and fourth-grade boys in their reading comprehension of material written with infrequent oral language patterns.

---

[12] Refer to page 4 of Chapter I for a list of the questions.

H6: Second-grade girls do not do significantly better than second-grade boys in their reading comprehension of material written with frequent oral language patterns.

H7: Second-grade girls do not do significantly better than second-grade boys in their reading comprehension of materials written with infrequent oral language patterns.

H8: Fourth-grade girls do not do significantly better than fourth-grade boys in their reading comprehension of material written with frequent oral language patterns.

H9: Fourth-grade girls do not do significantly better than fourth-grade boys in their reading comprehension of material written with infrequent oral language patterns.

The manner in which the hypotheses were tested is discussed below.

Hypothesis 1

By use of the chi square test, one can determine whether a set of observed frequencies is consistent with the frequencies that are expected if the hypothesis is accepted (Tate, 1955). If the discrepancy between the observed and expected frequencies is too large to be credited to sampling fluctuations, the hypothesis is rejected. With respect to the hypothesis under consideration, chi square was used to determine whether or not the number of subjects who contained a higher reading comprehension score on Test A (frequent oral language patterns) than Test B (infrequent oral language patterns) was significantly greater than the expected number.

Theoretical expectation was determined by dividing the number of subjects in each chi square analysis by two. This expectation of N/2 provided for a 50-50 chance that scores would be higher on one test than on another. If there were no discrepancy between this theoretical frequency and the observed frequency, then the value of chi square would be zero.

The statistic used was the following, where f_o = the observed frequency and f_e = the expected frequency:

$$x^2 = \frac{(f_o - f_e)^2}{f_e}$$

Hypotheses 2 and 3

For Hypotheses 2 and 3, the t test was used to test the significance of any observed difference in mean scores between the performance of second and fourth graders on the reading comprehension tests. Since the t test is based on the assumption of equal population variances, preliminary tests were carried out to provide a check on whether or not this was the case; the F ratio was computed for each hypothesis being tested. When the obtained value of F was significant at the .05 level, the following model for the t test was used (Winer, 1962):

$$t = \frac{(\bar{X}_a - \bar{X}_b) - (\mu_a - \mu_b)}{\sqrt{\frac{S_a^2}{N_a} + \frac{S_b^2}{N_b}}}$$

When the population variances could be assumed equal, this model for the t test was used (Hays, 1963):

$$t = \frac{(M_1 - M_2) \quad -E(M_1 - M_2)}{\sqrt{\frac{(N_1 S_1^2 + N_2 S^2{}_2)}{(N_1 + N_2 - 2)} \left(\frac{N_1 + N_2}{N_1 N_2} \right)}}$$

Hypotheses 4–9

For Hypotheses 4–9, the t test was used to test the significance of any observed difference in mean scores between the performance of boys and girls on the reading comprehension tests, both within and across grade level. To check the assumption of equal variance on which the t test is based, an F ratio was computed for each of the two groups being compared. If the obtained value of F was significant at the .05 level, the model for the t test which assumes unequal population variance was used as previously described for Hypotheses 2 and 3. When the population variances could be assumed equal, a different model for the t test, also described for Hypotheses 2 and 3, was used.

Summary of Analysis

Hypothesis 1 was tested by means of chi square. For the remaining eight hypotheses, one of two models for the t test was used depending on whether or not the population variances could be assumed equal.

Summary

In this chapter a general overview of the study preceded a discussion of how the investigator devised the instruments to measure children's reading comprehension of select oral language patterns. Procedures for controlling variables like content and grammatical complexity from test to test were outlined, and criteria for selecting the final items were presented. The chapter also contained information about the subjects, who were characterized as "average" in terms of IQ and socioeconomic background. Testing procedures were discussed and the statistics used to test the nine hypotheses were outlined.

III

RESULTS AND DISCUSSION

In the present study the relationship between reading comprehension and material written with frequent and infrequent oral language patterns was investigated in Grades 2 and 4. Three basic questions were stated:

Q1: Do significantly more second and fourth graders comprehend material written with frequent oral language patterns better than material written with infrequent oral language patterns?

Q2: Do fourth graders comprehend material written with frequent and infrequent oral language patterns significantly better than second graders?

Q3: Do girls comprehend material written with frequent and infrequent oral language patterns significantly better than boys?

Results relevant to the nine hypotheses are grouped according to the three questions.

The instruments used to measure reading comprehension of select oral language patterns were devised by the investigator for the present study. Because the tests were revised after preliminary piloting, additional estimates of reliability and validity were obtained after final data collection. Reliability and validity are discussed in the present chapter so that the results of the study can be considered along with this information about the measuring instruments.

RESULTS OF THE TESTS OF HYPOTHESES

Results Related to Reading Comprehension of Select Oral Language Patterns

In this section the results relating to Question 1, which concerned the differential effect of frequent and infrequent oral language patterns on reading comprehension, are presented.

> H1: The number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is better than reading comprehension of material written with infrequent oral language patterns is not significantly greater than the number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is not better than reading comprehension of material written with infrequent oral language patterns.

For the first chi square analysis, scores on Test A (frequent oral language patterns) and Test B (infrequent oral language patterns) were compared for all 300 subjects to see whether the number of higher scores on the former was significant. The results of the chi square analysis, with 1 degree of freedom and $\alpha = .05$, are presented in Table 14. As shown in the table, more subjects attained a higher reading comprehension score on Test A than on Test B (192 subjects who did, compared to 108 who did not). The number was significant at the .001 level; the null hypothesis was rejected. See Appendix D for the subjects' scores on Test A and Test B.

For further clarification of this finding, Table 15 was constructed to show the mean score, range, variance, and standard deviation for all the subjects on Test A and Test B. As indicated, the mean score on Test A was higher than on Test B (31.37 for Test A and 29.84 for Test B). The range of scores on both tests was the same and the variance was similar,

Table 15

Mean, Range, Variance and Standard Deviation for All 300 Subjects on Test A and Test B

Reading Comprehension Test	Mean	Range	Variance	Standard Deviation
Test A (frequent oral language patterns	31.37	10–37	18.90	4.35
Test B (infrequent oral language patterns)	29.84	10–37	20.94	4.58

although scores on Test B varied slightly more than scores on Test A (the variance was 20.94 on Test B and 18.90 on Test A). Frequency distributions of total scores on Test A and Test B for all subjects are presented in Appendix E.

Because it was felt that considering the total group might have concealed different results within some of the smaller groups of interest in the present study, chi square analyses were performed separately on test scores of subjects who were grouped by sex and by grade. The six additional analyses involved treating (1) each grade as a separate group, (2) each sex within grade as a separate group, and (3) each sex across grades as a separate group. The respective findings are presented below.

Each grade as a separate group. The result of the chi square analysis for the second graders is summarized in Table 14, which shows that the chi square value of 15.86 was significant at the .01 level. More subjects obtained their higher score on Test A than Test B (107 subjects who did; 56 who did not).

The result of the chi square analysis for the fourth grade is presented in Table 14. As shown, the greater number of subjects obtained their higher score on Test A than Test B (85 subjects who did; 52 who did not). The chi square value was significant at the .01 level.

Each sex as a separate group within Grades 2 and 4. Within the second grade, test scores were treated separately for girls and boys. The findings for the former are presented in Table 14. The chi square value of 6.47, with 1 degree of freedom, was significant at the .01 level. More subjects obtained their higher score on Test A than Test B (52 subjects who did; 29 who did not).

In Table 14, the results of the chi square analysis are presented for second grade boys. The table shows that the chi square value of 9.56, with 1 degree of freedom, was significant at the .01 level. A greater number of subjects scored higher on Test A than Test B (55 subjects who scored higher on Test A; 27 who did not).

Test scores were also analyzed separately for girls and boys within the fourth grade. Table 14 summarizes the findings for the fourth grade girls. The chi square value of 7.57, with 1 degree of freedom, reached the .01 level of significance. More subjects obtained their higher score on Test A than Test B (46 subjects who did; 23 who did not).

The result of the chi square analysis for fourth grade boys is presented in Table 14. As shown in the table, more boys obtained their higher score on Test A than Test B (39 subjects who did; 29 who did not), but the chi square value of 1.47, with 1 degree of freedom, was not significant at the .05 level.

Sex across Grades 2 and 4. The second- and fourth-grade girls and the second- and fourth-grade boys formed two groups which were treated separately. Table 14 summarizes the results of the chi square analysis of the girls' scores. As shown in the table, more girls obtained their higher score on Test A than Test B (98 subjects who did; 52 who did not). The chi square value of 15.36, with 1 degree of freedom, was significant at the .001 level.

More of the second- and fourth-grade boys obtained their higher score on Test A than Test B (94 subjects who did; 56 subjects who did not). As indicated in Table 14, the chi square value of 9.62, with 1 degree of freedom, was significant at the .01 level.

Results Related to Grade Level

In this section, results are presented for the two hypotheses that were derived from Question 2, which focused on the relationship between grade level and reading comprehension of material written with frequent and infrequent oral language patterns. In each case, the difference between the mean reading comprehension test scores of the two groups being compared was tested for significance; the t test was used for this purpose. In accordance with the requirements of the null hypothesis, the two-tailed t test was used. The .05 level of significance was accepted. A null hypothesis was rejected, consequently, if

$$t > t_{1 - (\alpha/2)}[N_1 + N_2 - 2] \text{ or if } t < t_{\alpha/2}[N_1 + N_2 - 2]$$

Because the variance of scores on both reading comprehension tests differed from grade to grade, F ratios were computed to test the homogeneity of variance. The F ratio was significant at the .05 level for grade level comparisons on each reading comprehension test; the model for the t test which is appropriate for unequal variance was used.

> H2: Fourth graders do not do significantly better than second graders in their reading comprehension of material written with frequent oral language patterns.

The results of the t test for second and fourth graders are presented in Table 16 for Test A (frequent oral language patterns). As shown in the table, the difference between the second- and fourth-grade means on Test A was significant at the .01 level (a mean of 32.98 for Grade 4 and 30.02 for Grade 2). The null hypothesis was consequently rejected. Frequency distributions of total scores on Test A are presented in Appendix E for both grades.

> H3: Fourth graders do not do significantly better than second graders in their reading comprehension of material written with infrequent oral language patterns.

Table 16 summarizes the results of the t test for second and fourth graders on Test B (infrequent oral language patterns). The difference in the second- and fourth-grade means was significant at the .01 level (a mean of 31.63 for Grade 4 and 28.34 for Grade 2). The null hypothesis was rejected. In Appendix E, frequency distributions of total test scores on Test B are presented for each grade.

Table 17 was constructed to elaborate the findings presented for Hypotheses 2 and 3. In the table, the mean, range, variance, and standard deviation of scores on both tests are compared for the two grades. It is clear that in addition to the lower mean score, the second graders had a greater range, variance, and standard deviation on Test A and Test B.

Results Related to Sex

In this section, results are discussed for the six hypotheses that were derived from Question 3, which dealt with the relationship of sex and reading comprehension of material written with frequent and infrequent oral language patterns. For each comparison of the mean scores, the t test models that assume either equal or unequal population variance were

Table 16

Results of the t test of Mean Scores for Second and Fourth Graders

Reading Comprehension Test	N	Second Grade		N	Fourth Grade		Mean Difference	t-value	Significance Level	Hypothesis
		Mean	SD		Mean	SD				
Test A (frequent oral language patterns)	163	30.02	4.96	137	32.98	2.72	2.96	6.58	<.01	H2 rejected
Test B (infrequent oral language patterns)	163	28.34	5.06	137	31.63	3.10	2.29	4.87	<.01	H3 rejected

Table 17

A Comparison of the Mean, Range, Variance and Standard Deviation of Scores for Grades 2 and 4

Reading Comprehension Test	Grade Level	Mean	Range	Variance	Standard Deviation
Test A					
(frequent oral	2	30.02	10–37	24.65	4.96
language patterns)	4	32.98	19–37	7.39	2.72
Test B					
(infrequent oral	2	28.34	11–36	25.62	5.06
language patterns)	4	31.63	16–37	9.59	3.10

used depending on whether or not the computed *F* ratio revealed significant differences in the variances of boys and girls.

In accordance with the requirement of the null hypothesis, the two-tailed *t* test was used. The .05 level of significance was accepted. A null hypothesis was rejected if

$$t > t_{1-(\alpha/2)} [N_1 + N_2 - 2] \text{ or if } t < t_{\alpha/2} [N_1 + N_2 - 2] .$$

> H4: Second and fourth grade girls do not do significantly better than second and fourth grade boys in their reading comprehension of material written with frequent oral language patterns.

The variances for the girls and boys differed significantly at the .05 level according to the *F* ratio. Thus the *t* test model, appropriate for unequal variance was used. Results of the *t* test are presented in Table 18. As shown in the table, the difference between the mean scores of the girls and boys on Test A (frequent oral language patterns) was significant at the .02 level. The higher mean favored the girls (31.97 for the girls and 30.77 for the boys). The null hypothesis was consequently rejected.

Table 18

Results of the t Test of Mean Scores for Boys and Girls Across Grade Levels

Reading Comprehension Test	Girls			Boys			Mean Difference	t-value	Significance Level	Hypothesis
	N	Mean	SD	N	Mean	SD				
Test A (frequent oral language patterns)	150	31.97	3.77	150	30.77	4.79	1.21	2.46	<.02	H4 rejected
Test B (infrequent oral language patterns)	150	30.27	4.29	150	29.41	4.82	.85	1.61		H5 accepted

H5: Second- and fourth-grade girls do not do significantly better than second- and fourth-grade boys in their reading comprehension of material written with infrequent oral language patterns.

According to the computed *F* ratio, the variances of the girls and boys did not differ significantly at the .05 level. The *t* test model which assumes equal variances was used. As shown in Table 18, the difference between the mean scores of girls and boys on Test B (infrequent oral language patterns) was not significant at the .05 level, but the higher mean was obtained by the girls (30.27 for the girls and 29.41 for the boys). The null hypothesis was accepted.

H6: Second-grade girls do not do significantly better than second-grade boys in their reading comprehension of material written with frequent oral language patterns.

To test this hypothesis, the *t* test model which assumes unequal variance was used because the *F* ratio was significant at the .05 level. The mean scores for Test A (frequent oral language patterns) indicated that second grade girls scored higher than second grade boys (30.46 for the girls and 29.59 for the boys). As shown in Table 19, however, this difference was not significant at the .05 level. The null hypothesis was accepted.

H7: Second-grade girls do not do significantly better than second-grade boys in their reading comprehension of material written with infrequent oral language patterns.

Since the *F* ratio of the two variances was not significant at the .05 level, the *t* test model which assumes equal variances was used. Results of the *t* test for second grade boys and girls on Test B (infrequent oral language patterns) are presented in Table 19. Although the higher mean was attained by the girls (28.75 for the girls and 27.93 for the boys), the difference was not significant at the .05 level. The null hypothesis was accepted.

Table 19

Results of the t Test of Mean Scores for Boys and Girls in Second Grade

Reading Comprehension Test		Girls			Boys		Mean Differ-ence	t-value	Signifi-cance Level	Hypothesis
	N	Mean	SD	N	Mean	SD				
Test A (frequent oral language patterns)	81	30.46	4.23	82	29.59	5.59	.87	1.15		H6 accepted
Test B (infrequent oral language patterns)	81	28.75	4.74	82	27.93	5.35	.83	1.03		H7 accepted

> H8: Fourth-grade girls do not do significantly better than fourth-grade boys in their reading comprehension of material written with frequent oral language patterns.

The F ratio of the two variances was significant at the .05 level; the t test model which assumes unequal variances was used. As shown in Table 20, the mean scores for fourth-grade girls and boys were not significantly different at the .05 level. However, the higher mean score was obtained by the girls (33.75 for the girls and 32.19 for the boys). The null hypothesis was accepted.

> H9: Fourth-grade girls do not do significantly better than fourth-grade boys in their reading comprehension of material written with infrequent oral language patterns.

Because the F ratio of the variances was significant at the .05 level, the t test model which assumes unequal variance was used. Results of the t test for fourth-grade girls and boys on Test B (infrequent oral language patterns) are summarized in Table 20. The table indicates that although the higher mean score was attained by the fourth grade girls (32.04 for the girls and 31.21 for the boys), the difference was not significant at the .05 level. The null hypothesis was accepted.

To elaborate the findings for Hypotheses 4—9, which dealt with the relation of sex to reading comprehension of select oral language patterns, Table 21 was constructed. In the table, the mean, range, variance, and standard deviation for Test A and Test B are summarized for girls and boys within and across grade levels. From inspection of the table it is clear that in addition to obtaining consistently lower mean scores than the girls, the boys were characterized by greater variance and a greater range of scores. Two exceptions to the latter were the total group of boys and the second-grade boys who, in Test B, had a smaller range of scores than the corresponding group of girls.

Table 20

Results of the *t* Test of Mean Scores for Boys and Girls in Fourth Grade

Reading Comprehension Test	Girls			Boys			Mean Differ-ence	t-value	Signifi-cance Level	Hypothesis
	N	Mean	SD	N	Mean	SD				
Test A (frequent oral language patterns)	69	33.75	2.04	68	32.19	3.09	.56	1.25		H8 accepted
Test B (infrequent oral language patterns)	69	32.04	2.81	68	31.21	3.33	.84	1.60		H9 accepted

Table 21

Mean, Range, Variance, and Standard Deviation of Scores on Test A and Test B for Girls and Boys Within and Across Grades Two and Four

Reading Comprehension Test	Subjects	N	Mean	Range	Variance	Standard Deviation
Test A (frequent oral language patterns)	All girls	150	31.97	14–37	14.20	3.77
	All boys	150	30.77	10–37	22.99	4.79
	Second-grade girls	81	30.46	14–36	17.85	4.23
	Second-grade boys	82	29.59	10–37	31.28	5.59
	Fourth-grade girls	69	33.75	28–37	4.16	2.04
	Fourth-grade boys	68	32.19	21–37	9.58	3.09
Test B (infrequent oral language patterns)	All girls	150	30.27	11–37	18.40	4.29
	All boys	150	29.41	13–36	23.25	4.82
	Second-grade girls	81	28.75	11–37	22.51	4.74
	Second-grade boys	82	27.93	13–34	28.66	5.35
	Fourth-grade girls	69	32.04	21–37	7.90	2.81
	Fourth-grade boys	68	31.21	17–36	11.09	3.33

Summary of Results

The results for each of the nine hypotheses are summarized as follows:

Hypothesis 1. In general, significantly more second and fourth graders obtained higher reading comprehension scores on material written with frequent oral language patterns (Test A) than on material written with infrequent oral language patterns (Test B). The exception was fourth-grade boys, whose higher scores on Test A did not reach an acceptable level of significance.

Hypothesis 2. Fourth graders comprehended material written with frequent oral language patterns (Test A) significantly better than second graders.

Hypothesis 3. Fourth graders comprehended material written with infrequent oral language patterns (Test B) significantly better than second graders.

Hypothesis 4. Second- and fourth-grade girls comprehended material written with frequent oral language patterns (Test A) significantly better than second- and fourth-grade boys.

Hypothesis 5. Second- and fourth-grade girls did not do significantly better than second- and fourth-grade boys in their reading comprehension of material written with infrequent oral language patterns (Test B).

Hypothesis 6. Second-grade girls did not do significantly better than second-grade boys in their reading comprehension of material written with frequent oral language patterns (Test A).

Hypothesis 7. Second-grade girls did not do significantly better than second-grade boys in their reading comprehension of material written with infrequent oral language patterns (Test B).

Hypothesis 8. Fourth-grade girls did not do significantly better than fourth-grade boys in their reading comprehension of material written with frequent oral language patterns (Test A).

Hypothesis 9. Fourth-grade girls did not do significantly better than fourth-grade boys in their reading comprehension of material written with infrequent oral language patterns (Test B).

A summary of the results of the tests of hypotheses is presented in Table 22.

DISCUSSION OF THE RESULTS

A discussion of the results follows for the hypothesis related to reading comprehension of select oral language patterns, the hypotheses related to grade level, and the hypotheses related to sex.

Results Related to Reading Comprehension of Select Oral Language Patterns

Rejection of the first hypothesis was expected on the basis of Ruddell's research (1963). With the results of Hypothesis 1, Ruddell's major finding was confirmed, elaborated, and extended in three ways: (1) applicability of the finding to children at the primary as well as the intermediate level of elementary school was demonstrated; (2) the finding was based on measuring instruments which did not systematically bias the performance of children on the infrequent oral language patterns;[1] (3) the

---

[1] As previously discussed in the review of Ruddell's study (1963) in Chapter I, it was suggested that his use of the cloze procedure may have biased his results.

Table 22

Summary of the Results of the Tests of Hypotheses

Statistic Used	Group Considered	df	Signifi-cance Level	Hypoth-esis	Decision
x^2	All subjects	1	<.001	H1	reject
	Second graders	1	<.001		
	Fourth graders	1	<.01		
	Second-grade girls	1	<.05		
	Second-grade boys	1	<.01		
	Fourth-grade girls	1	<.01		
	Fourth-grade boys	1	----		
	All girls	1	<.001		
	All boys	1	<.01		
t test	Second and fourth graders (Test A)	298	<.01	H2	reject
	Second and fourth graders (Test B)	298	<.01	H3	reject
	All girls and all boys (Test A)	298	<.05	H4	reject
	All girls and all boys (Test B)	298	----	H5	accept
	Second-grade girls and boys (Test A)	161	----	H6	accept
	Second-grade girls and boys (Test B)	161	----	H7	accept
	Fourth-grade girls and boys (Test A)	135	----	H8	accept
	Fourth-grade girls and boys (Test B)	135	----	H9	accept

results were observed for a sample of children from a different part of the country than the samples used in Ruddell's and Strickland's studies (both were conducted in the elementary schools of Bloomington, Indiana). The applicability of Strickland's findings (1962) to children other than those from Bloomington, Indiana, is thus strongly suggested.

There is nothing in the literature to explain the performance of fourth-grade boys, whose higher scores on Test A did not reach an acceptable level of significance. Indeed, from Ruddell's (1963) statement that his fourth-grade boys had a disproportionate amount of difficulty reading infrequent oral language patterns, one would have expected significantly more boys to obtain their higher score on Test A (frequent oral language patterns) than on Test B (infrequent oral language patterns). Two possible explanations for this are suggested and discussed.

1. The material in Test B was more appealing than in Test A, so performance on the former was better regardless of language patterns.

As indicated in Chapter II, the two tests were rigorously controlled to assure similarity in sentence and picture content, vocabulary, and grammatical complexities. This explanation is not logical in terms of the constraints imposed on the materials.

2. There was a relatively larger portion of good readers among the boys than among the girls; good readers would not be expected to differentiate their performance on the two types of language patterns to the same extent as poor readers (see a discussion of test validity at the end of the chapter for elaboration of this point).

To see whether this was the case, the reading comprehension grade placement scores on the Stanford subtest for fourth grade girls and boys were rated in one of three categories that were based on the corresponding stanines: 1—3 (corresponding to the 2.3—3.8 grade placement scores = poor readers; 4—6 (corresponding to the 3.9—5.9 grade placement scores)

= average readers; 7–9 (corresponding to the 6.0–9.5 grade placement scores) = good readers. The obtained frequencies in each category are presented in Table 23 for girls and boys. As indicated, the scores for both sexes were similarly rated; there is no substantial evidence to support the second explanation. Clearly more research is needed to clarify the performance of fourth-grade boys in the present study.

Table 23

The Number of Poor, Average, and Good Readers Among Fourth-Grade Girls and Boys as Determined by Scores from the Paragraph Meaning Section of the Stanford Achievement Test

	Poor Readers (Stanine 1-3)	Average Readers (Stanine 4-6)	Good Readers (Stanine 7-9)
Girls	7	38	24
Boys	6	35	27

The Results Related to Grade Level

Significant differences between the performance of second and fourth graders on both reading comprehension tests were expected. Possible explanations for these differences are listed and discussed below.

1. Fourth graders read better than second graders.

This is not a profound nor contested observation about reading performance of children in successively higher grade levels; it is based on the fact that reading is a skill that improves with practice and experience. An interesting problem is raised by the tests used in the present study, however, because they were not primarily intended to measure skills that are taught in reading programs at the elementary school. Typical measures of reading ability such as vocabulary and reading speed were carefully controlled in the materials devised for the study so that they would not be heavily weighted factors in the children's reading comprehension scores.

As discussed more fully in Chapter II, these additional controls were imposed to limit the effect of distinctly different levels of reading achievement between the two grades prior to testing: (1) no time limit so that second graders would not be penalized for taking longer to read through the tests; (2) practice items to equalize the subjects' experience with the particular nature of the reading comprehension task; (3) standardized, clear directions for all subjects; (4) a reasonable test length which, in piloting, was shown not to overtax the endurance limitations of the younger subjects; (5) the use of words which are presumed relatively easy for primary grade children (predominantly from the Stone List); (6) directions for children to ask for help with any word they did not know; (7) the use of a measure of reading comprehension that was assumed to be relatively new and interesting to subjects of both grades.

If it can be assumed that differences between the grades in terms of decoding ability were minimized rather than maximized by the controls imposed on the materials, then the first explanation—fourth graders read better than second graders—is rejected as it is presently worded and a second explanation is proposed as a more acceptable alternative:

2. Second graders are less successful than fourth graders in comprehending the relationship between oral and written language patterns.

From the greater variability of second-grade scores on both Test A and Test B, it is apparent that children at this level of reading ability and experience have not yet developed as much flexibility in handling diverse language patterns in written material as fourth graders. Reasons for this are beyond the scope of the present study, but some suggestions are possible from two areas: (1) a linguistic analysis of the reading process, and (2) research on children's written language structures.

With respect to the first area, Lefevre (1964) has emphasized the need to read for intonation patterns in sentences—to comprehend the "melodies of the printed page." From the very beginning of reading instruction, therefore, a child must learn to read sentences; beginning with isolated words results in the unnatural practice of giving a full stress to each word—the opposite of what is done in speaking. An early understanding of the relation between spoken and written language is consequently hampered. From this point of view, the different performance of second and fourth graders might have been due, in part, to the still incomplete understanding by the former of (1) the relation between spoken and written language, and (2) the constraints of written language that need to be understood and compensated for, examples being spatial rather than temporal orientation of language, and missing intonation clues.

From the area of research on children's written language structures, Hunt (1965) has concluded that fourth graders understand complex grammatical structures because they write them often and accurately. The fact that these structures were not produced as frequently or with as much variety by fourth graders as by the older children in his study, however, led Hunt to recommend that a program should be developed to widen the younger children's span of grammatical attention and concern. The intent would be to work toward performance that is comparable to that of twelfth graders.

The implied difference in grammatical awareness between the grades in Hunt's study is applicable to the distinction that can be made between the grades studied in this investigation. Compared to fourth graders, second graders might not be as consciously aware of how word order in the visual context of reading relates to meaning even though they are capable of producing a variety of complex word orders in their oral language.

Indeed, second graders might be so busy looking at and considering each word that they miss the meaning of whole language patterns.

3. The materials were too easy for fourth graders.

This explanation could be possible for one of two reasons: (1) either the materials were too easy because they were written at the second-grade level, or (2) the materials were too easy because the patterns themselves were too easy for children at the fourth-grade level. With respect to the latter, it is suggested that the tests were not as reliable measures of reading comprehension for fourth graders because they tested something, defined as language pattern difficulty, which was not as relevant a concern for this grade as for second grade.[2] Perhaps the infrequent oral language patterns selected for Test B were not infrequent enough for the older subjects; selection of less frequent patterns might have increased the difficulty of the items in Test B.

Acceptance of either of the two alternatives suggested above, however, does not affect the fact that despite the relative ease with which fourth graders handled the materials, a significantly greater number of them obtained higher scores on Test A than on Test B. The implications of the results of the first hypothesis, which stated that the number of second and fourth graders who obtained higher scores on Test A than Test B would not be significant, should not be restricted to second graders.

Results Related to Sex

The general finding that boys did less well on each test than girls was expected on the basis of many researchers who have found differential performance of the sexes in elementary school reading and language related tasks. This difference due to sex has been noted from the very beginning stages of reading readiness, where girls have generally been found superior (Barrett, 1962). Broader studies of the relation of sex to performance in several language tasks have generally found differences, though not all of them have been consistent or significant ones.[3]

---

[2] See the discussion of test reliability in this chapter and the results of item analysis by grade level, found in Appendix F, for further clarification of this possibility.

Loban's (1963) finding that boys who performed poorly in language obtained the lowest reading scores while boys who did well in language obtained among the highest reading scores was supported, with few exceptions, by the results of the six hypotheses related to sex in the present study. The poorest scores for all 300 subjects was attained by a boy (10, on Test A), and the highest possible score on each test, 37, was also attained by a boy. These results emphasize the need to consider great variation of ability within each sex rather than treating members of one sex as a homogeneous group.

Reliability of the Reading Comprehension Tests

A broad definition of test reliability includes reference to these three basic characteristics of a test (Cronbach, 1960): (1) the extent to which the measuring instrument produces similar results after repeated use (stability and dependability); (2) the extent to which errors of measurement are present (random chance or error); and (3) the extent to which the given results are a true measure of what is being assessed (accuracy). In accordance with this definition, three different measures of reliability were obtained for each reading comprehension test; the results of each are discussed below.

---

[3] McCarthy (1954) found girls superior to boys in most aspects of language, but Templin (1957), Strickland (1962), Loban (1963), and Riling (1965) did not find clear-cut, consistent sex differences.

First, reliability was examined in terms of stability. Due to the limited scope of the study and restrictions of time, the usual test-retest method was not feasible. Instead, stability across rather than within tests was evaluated by correlating scores on Test A and Test B to see to what extent good or poor performance on one was indicative of good or poor performance on the other. For this purpose, the Pearson product-moment coefficient of correlation was determined. This coefficient measured the amount of relationship between each set of scores on Test A and Test B; it was based on the products of pairs of these scores.[4] The obtained coefficient of correlation was .83, which was significant at the .01 level. Therefore, performance on Test A was significantly related to performance on Test B.

Second, reliability was assessed in terms of the standard error of measurement, which provides an indication of the reliability of individual scores (Lindvall, 1967). Unlike the reliability coefficient, the standard error of a score is independent of the variability of the group and consequently is not affected by group homogeneity, a characteristic that appreciably lowers the reliability coefficient (Bloom, 1966). The standard errors for Test A and Test B are summarized in Table 24 both within and across

---

[4] The basic formula for the Pearson product-moment correlation is

$$r_{xy} = \frac{\Sigma xy}{\sqrt{(\Sigma x^2)\ (\Sigma y^2)}} \quad \text{where } x = (X - \overline{X}),\ y = (Y - \overline{Y}), \text{ and x and}$$

y are the two variables being considered (Tate, 1955, p. 238).

Table 24

Hoyt Reliability Coefficients Within and Across Grades for Test A and Test B

Test	Grade Level					
	Second		Fourth		Second & Fourth	
	R	SE	R	SE	R	SE
Test A (frequent oral language patterns)	.81	2.14	.58	1.74	.79	1.97
Test B (infrequent oral language patterns)	.81	2.20	.59	1.95	.78	2.10

Grades 2 and 4. Inspection of the table reveals that (1) the standard errors for Test B within and across grades were somewhat higher than for Test A, and (2) the standard errors for the second grade were somewhat higher on each test than those for the fourth grade. Since the maximum standard error was 2.20, the tests were considered fairly reliable measures of individual scores.

Third, the degree of internal consistency, or accuracy, was determined by using the Hoyt Reliability Coefficient.[5] The reliability coefficients for each test within and across Grades 2 and 4 are summarized in Table 24. On the basis of information presented in the table, these observations can be made: (1) Within each grade, reliability coefficients of Test A and Test B were similar; (2) the reliability coefficients of both tests were considerably higher for second-grade subjects than fourth-grade subjects.

---

[5] The Hoyt Reliability Coefficient was obtained from a 1604 computer program, FORTAP (Baker and Martin, 1968).

The low test reliability for the fourth grade[6] can be examined in light of the findings for the tests of hypotheses, discussed in the preceding section of this chapter. From the results related to grade level differences in test performance (H2 and H3), it is obvious in Table 17 that the variance of scores for the second grade was almost three times greater on Test A than for fourth grade, and more than three times greater for the second grade on Test B. Since a reliability coefficient depends on the variability of scores (Anastasi, 1960), lack of variability at the fourth grade level probably contributed heavily to the low reliability coefficient of each test for fourth grade subjects.

From the means and ranges of scores for the fourth grade, information also presented in Table 17, it was apparent that scores were not distributed at both ends of the scale but were, instead, clustered near the top (see Appendix F for the frequency distribution of scores for the fourth grade). Tests as negatively skewed as these for fourth graders may have been satisfactory for measuring differences at the lower end of the group of subjects, but not for measuring differences at the upper end; the test could not distinguish between scores of subjects who tied at or near the perfect score of 37.

Other factors that can affect test reliability such as the testing environment, instructions, and time limit (Anastasi, 1961) were controlled across grades in that (1) all subjects were tested within their own classroom, a familiar environment, (2) the instructions were standardized, and (3) there was no time limit. These factors probably did not contribute to the discrepancy in reliability from grade to grade.

---

[6] Bloom (1966) has stated that a reliability coefficient of .60 is adequate for group testing.

In conclusion, the considerably smaller variance of scores obtained on both Test A and Test B by fourth graders probably accounted in large part for the low reliability coefficients at the fourth grade level. Explanations for the smaller variance and possible implications were presented in the discussion of the results that preceded this section of the chapter. See Appendix F for results of an item analysis of Test A and Test B in terms of item reliability (R) and difficulty (P) for each grade.

Summary

Results of three measures of reliability were discussed for Test A and Test B. On the basis of these results it seemed that the tests were fairly reliable measures of individual scores. The tests were accepted as reliable instruments for the second-grade subjects but considerably less reliable instruments for the fourth-grade subjects. The smaller variance of scores for the latter probably contributed substantially to the lower reliability.

Validity of the Reading Comprehension Tests

Two types of validity were considered in devising and evaluating the two reading comprehension tests. They were construct validity, which has to do with the meaning of test scores in terms of psychological concepts and theory (Cronbach, 1960), and content validity, which has to do with the appropriateness of the content of the test items in terms of what the test as a whole is measuring (Lindvall, 1967). In this section each of the two kinds of validity is discussed.

Construct Validity

Theories of language learning and its relation to reading achievement have not yet been tested thoroughly. Researchers have made interesting but inconclusive reference to the fact that children who are good readers have considerable facility with spoken language; these children use more complex language structures and they use the basic structures with flexibility (Strickland, 1962; Loban, 1963). Conversely, children who are poor readers seem to lack the kind of flexibility with patterns of language structure that characterizes mature language.

Learning theorists and educators generally support the idea of moving from the familiar to the unfamiliar in teaching children.[7] With respect to reading, a task which involves learning to obtain meaning from written language patterns, it is also logical to move from the familiar to the unfamiliar. In a sense, this is already done. Rigid control over vocabulary is typically imposed in basal readers; a little vocabulary is developed over a long period of time. In terms of the context of these words—the sentence or language pattern—however, such control is not apparent (Strickland, 1962; Riling, 1965).

It is reasonable to suggest that learning the relation between written and spoken language would be facilitated greatly if children began reading language patterns that were familiar to them before they moved, gradually, to less familiar patterns. Recent research on children's oral language provides a reasonable way to decide whether or not certain patterns are familiar or unfamiliar to children: decisions can be based on observations of oral language patterns that appear very frequently or very infrequently in their speech. Thus, it is suggested that the terms "familiar" and "unfamiliar" be made synonymous, respectively, with "frequent" and "infrequent" patterns as they appear in children's oral language.

---

[7] Lindgren (1962) provides a typical example of this idea in his discussion of the conceptual approach to learning: ". . . learning is reinforced when the various experiences of life . . . are seen by the students as interrelated. Information and skills are more readily learned and retained if we see how they are related to other things we know" (p. 267).

In this manner, then, the theoretical framework for the present study was derived from two sources: (1) learning theory which emphasizes moving from the familiar to the unfamiliar for more effective and efficient learning, and (2) the assumption that one can determine language patterns for written material which are familiar or unfamiliar to children by observing which patterns children use frequently or infrequently in their oral language. From this framework, a major hypothesis of the study was stated:

> H1: The number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is better than reading comprehension of material written with infrequent oral language patterns is not significantly greater than the number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is not better than reading comprehension of material written with infrequent oral language patterns.

Support for the reason to suspect differences in reading comprehension of frequent (familiar) and infrequent (unfamiliar) patterns was derived from basically two sources. First, from an observation of research in this area (Ruddell, 1963; Tatham, 1967; Tatham, 1968), it was clear that there is unanimous though limited evidence that children comprehend material written with frequent oral language patterns significantly better than material written with infrequent oral language patterns.[8] Second, from their analysis of children's oral language, some researchers have indicated that certain patterns are the building blocks of children's oral language; they are used frequently regardless of age or intelligence. Since these patterns are common, one would expect children to have the least difficulty reading them. On the other hand, since the use of infrequent patterns is not common to all children, one would expect that the ability to read them would distinguish between levels of achievement more than the ability to read the "easier" frequent patterns. A logical hypothesis that emerges from this line of reasoning is that scores on Test B (infrequent patterns) should be a better predictor of general reading comprehension, as measured by the Stanford subtest, than scores on Test A (frequent patterns).

---

[8] Fourth graders were the subjects for each of the three studies.

To determine whether or not this was the case, a multiple regression analysis was performed by means of REGANI.[9] For the analysis, the two independent variables were scores on Test A and Test B; the dependent variable was scores from the Stanford subtest. As shown in Table 25, the partial regression coefficients of both Test A and Test B were significant at the .01 level. In other words, both independent variables contributed significantly to the scores on the Stanford subtest. To look at this another way, partial correlations of the Stanford subtest with each of the reading comprehension tests were determined as part of the REGANI program. The partial correlations were quite low, as indicated in Table 26. This means that Test A and Test B apparently had a great deal in common. From its higher partial correlation with the Stanford subtest, however, Test B provided evidence of an element that was unique to it alone. This unique element, probably difficulty, correlated more highly with the Stanford subtest than Test A.

---

[9] Susanta Guha (1966). Given N number of observations on p variables, REGANI can be used to solve linear regression equations by the method of least squares.

Table 25

Partial Regression Coefficients of Test A and Test B

Variable	Partial Regression Coefficients	SE	_t_-value
Stanford subtest	-32.76	6.19	-5.29
Test A	1.01	.35	2.88**
Test B	1.44	.34	4.26***

** $p < .01$
*** $p < .001$

Table 26

Partial Correlations of Test A and Test B with the Paragraph Meaning Section of the Stanford Achievement Test

Test	Partial Correlation
Test A	.17**
Test B	.25**

** $p < .01$

In conclusion, from both the partial correlations and the multiple regression analysis it would seem that performance on Test B rather than Test A was a somewhat better indicator of general reading comprehension as measured by the Stanford subtest.

A last statistical measure of validity was obtained by determining the coefficient of correlation between Test A, Test B, the combined scores, and the Stanford subtest. The resulting correlation matrix is summarized in Table 27. As noted in the table, all the correlations were significant at the .01 level.

Table 27

Validity Coefficient of Correlation between Test A, Test B, the Combined Scores, and the Paragraph Meaning Section of the Stanford Achievement Test

	1	2	3	4
Test A	1.00	.83**	.95**	.56**
Test B		1.00	.95**	.58**
Combined Scores (A + B)			1.00	.61**
Stanford Subtest				1.00

** $p < .01$

The high intercorrelation of Test A with Test B (.83) could be interpreted in at least one of three ways (Cronbach, 1960): (1) reading comprehension of frequent oral language patterns (Test A) caused or influenced reading comprehension of infrequent patterns (Test B); (2) reading comprehension of infrequent oral language patterns (Test B) caused or influenced reading comprehension of frequent patterns (Test A); or (3) scores on both Test A and Test B were influenced by some common factor or factors. From inspection of the coefficients of correlation of Test A and Test B with the Stanford subtest, it is evident that general reading ability, defined by scores on the Stanford subtest, was a factor in determining performance on the materials written with frequent and infrequent oral language patterns.

Content Validity

As discussed in Chapter II, the frequent and infrequent oral language patterns were selected from among those that appeared in the oral language of both second and fourth graders, according to the findings of Strickland's study (1962). Since Strickland reported a total of over 200 different patterns, it was obvious that only a select number could be used in the kind of tests devised for the present study.

Two small samples from the total universe of children's oral language patterns were designated as "frequent" and "infrequent," and from within each sample, a smaller number of patterns was selected for use in the study. It was assumed that findings from the subsamples could be generalized to the samples and, finally, to the universe of patterns labeled "frequent" and "infrequent" in the oral language of children. Clearly, then, the patterns selected for each test were representative but by no means inclusive.

Within a language pattern there were several ways to fill in each slot or movable. As previously stated, a decision was made to include a variety of fillers such as adjectives, phrases, clauses, and compound nouns, verbs, or objects. Once again, the types of subordination and grammatical complexities that were used were representative but not inclusive of the possible ways of filling in each slot and movable.

As a final check on content validity, a judge who was thoroughly familiar with Strickland's work and similar studies of oral language patterns analyzed the 37 items in each reading comprehension test to be sure that the patterns used corresponded with the patterns that had been selected for use by the investigator.

Summary

Construct and content validity were discussed for the two measuring instruments devised for the study by the investigator. Both separately and together, the tests seemed to be based on valid, logical assumptions from learning theory and recent studies of children's oral language and its relationship to reading. In terms of content, the tests were representative of patterns designated "frequent" and "infrequent" in the oral language of second and fourth graders.

Summary

Results of the tests of the nine hypotheses of the study were presented and discussed in this chapter. In addition, obtained estimates of the reliability and validity of the two measuring instruments devised by the investigator were presented.

IV

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Background and Purpose

The relationship between reading comprehension and material written with frequent and infrequent oral language patterns of children was investigated in the present study.

Linguists have emphasized that the area of reading instruction must be concerned much more than it is with understanding language and how it is used in spoken and written messages (Carroll, 1964)—that is, how spoken and written language are related. For many children the process of learning this concept is made unnecessarily difficult by the "unnatural language" of their reading materials. Vocabulary is typically rigidly controlled in their texts, but there is no apparent evidence of control over the context of words—the language patterns which the children read (Strickland, 1962; Riling, 1965).

Writing materials for children which are structured more like the way they speak is a linguistically sound alternative to obstructing children's understanding of the fundamental relationship between the language they bring to school and the language they encounter in written material. In the single study which investigated this possibility, fourth graders comprehended material written with patterns children use frequently in their oral language significantly better than material written with patterns children use infrequently in their oral language (Ruddell, 1963). More research is needed to confirm this finding for different groups of children at different grade levels. Determining whether or not certain oral language patterns used by children are easy or difficult for them to comprehend in written material could bring a relevant but heretofore neglected characteristic of the reader—his oral language—to the concept of readability.

The present study was undertaken to investigate further the relationship between reading comprehension and material written with select oral language patterns at Grades two and four. The relation of sex to reading comprehension of these patterns was of additional interest.

More specifically, the study was designed to answer three basic questions:

Q1: Do significantly more second and fourth graders comprehend material written with frequent oral language patterns better than material written with infrequent oral language patterns?

Q2: Do fourth graders comprehend material written with frequent and infrequent oral language patterns significantly better than second graders?

Q3: Do second and fourth grade girls comprehend material written with frequent and infrequent oral language patterns significantly better than second and fourth grade boys?

Related Research

The review of the literature was divided into two sections: studies of elementary school children's oral language, and experimental studies using written material based on an analysis of children's oral language. The first section was limited to investigations from 1959-1968 that were based on a broad analysis of oral language obtained from relatively large samples of elementary school children. From the five studies conducted during this period, several conclusions were drawn: (1) the researchers demonstrated that children's language can be analyzed by methods derived from structural linguistics and transformational grammar and that information from these methods permits a variety of language characteristics to be described and compared for children of widely different abilities and age levels; (2) a relationship between reading comprehension and select oral language characteristics like the use of subordination and movables was suggested; (3) from some of the findings, it was apparent that children's materials are not written with patterns that children use in their oral language.

The second section of the review dealt with the only study in this area at present. Ruddell (1963) constructed materials written with frequent and infrequent oral language patterns of fourth graders, as determined by Strickland (1962), and he tested fourth graders' reading comprehension of them. He found that children comprehended materials written with the frequent patterns significantly better than materials that used the infrequent patterns. Ruddell concluded that reading comprehension is a function of the similarity of oral and written language.

Information from the five studies of children's oral language and the inquiry into children's reading comprehension of select oral language patterns provided both the framework and the impetus for the present investigation.

Method

To determine whether or not children comprehend material written with frequent oral language patterns better than material written with infrequent oral language patterns, two reading comprehension tests were devised by the investigator. Test A was written with patterns that appear frequently in the oral language of both second and fourth graders; Test B was written with language patterns that appear infrequently in the oral language of second and fourth graders.

The patterns were selected from Strickland's (1962) tabulation of the frequency with which oral language patterns of second and fourth graders appear; five highly frequent patterns were selected for Test A and six less frequently used ones were selected for Test B. They appeared in proportion to the number of times children use them in their oral language.

Certain restrictions were imposed on the materials to prevent bias at either grade and to assure similar content and complexity (other than the variable of pattern complexity) from test to test. The variables of vocabulary, content, and grammatical complexity were defined and then controlled within and across tests. Every effort was made to make the actual decoding process as easy as possible for the second graders.

Each test consisted of 37 items (sentences) which were selected after preliminary piloting according to (1) how well the distractors functioned, (2) how well the item correlated with the rest of the test, and (3) how difficult the item was. Three pictures were drawn for each sentence; one illustrated the exact content of the sentence while the other two pictures differed in one or more details. Reading comprehension was determined by the child's ability to read each sentence and select one of the three pictures that best represented the content of the sentence.

Subjects for the study were drawn from two elementary schools in Madison, Wisconsin. All classrooms of second and fourth graders in each school were used, yielding a final total of 163 second graders, 81 girls and 82 boys, and 137 fourth graders, 69 girls and 68 boys.

The subjects were tested by classroom for three separate testing sessions. During one week in the spring, the reading comprehension tests were administered on two consecutive mornings to each class. During the following week, the Paragraph Meaning section of the Stanford Achievement Test was administered. Subjects were randomly assigned to read either Test A or Test B first.

Derived from the three questions, nine hypotheses of the study were tested:

H1: The number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is better than reading comprehension of material written with infrequent oral language patterns is not significantly greater than the number of second and fourth graders whose reading comprehension of material written with frequent oral language patterns is not better than reading comprehension of material written with infrequent oral language patterns.

H2: Fourth graders do not do significantly better than second graders in their reading comprehension of material written with frequent oral language patterns.

H3: Fourth graders do not do significantly better than second graders in their reading comprehension of material written with infrequent oral language patterns.

H4: Second and fourth grade girls do not do significantly better than second and fourth grade boys in their reading comprehension of material written with frequent oral language patterns.

H5: Second and fourth grade girls do not do significantly better than second and fourth grade boys in their reading comprehension of material written with infrequent oral language patterns.

H6: Second grade girls do not do significantly better than second grade boys in their reading comprehension of material written with frequent oral language patterns.

H7: Second grade girls do not do significantly better than second grade boys in their reading comprehension of material written with infrequent oral language patterns.

H8: Fourth grade girls do not do significantly better than fourth grade boys in their reading comprehension of material written with frequent oral language patterns.

H9: Fourth grade girls do not do significantly better than fourth grade boys in their reading comprehension of material written with infrequent oral language patterns.

Hypothesis 1 was tested by chi square analysis. For the remaining eight hypotheses, one of two models for the $\underline{t}$ test was used depending on whether or not the population variances could be assumed equal.

Results

In Chapter III, the results of tests of the hypotheses were presented and discussed; measures of reliability and validity of the two instruments devised by the investigator were also presented.

The findings were grouped to provide answers to the three questions. With respect to the first question, the chi square analysis of Hypothesis 1 was reported, where 1 degree of freedom was used and the .05 level of significance was accepted. For the total group of 300 subjects the null hypothesis was rejected. For all the subgroups but one—the fourth-grade boys—a significantly larger number of subjects obtained their higher reading comprehension score on Test A (frequent oral language patterns) than on Test B (infrequent oral language patterns); subgroups were defined by grade and by sex within and across grades.

Derived from Question 2, Hypotheses 2 and 3 dealt with the relationship of grade level and the reading comprehension scores on Test A and Test B. $\underline{t}$ tests were used to determine whether or not obtained differences in mean scores were significant at the .05 level. Since significant differences were found, Hypotheses 2 and 3 were rejected.

Hypotheses 4—9 which related to Question 3, focused on the relationship of sex and reading comprehension scores on Test A and Test B. To test the differences between mean scores of boys and girls on the two tests, $\underline{t}$ tests were used. Based on results of the $\underline{t}$ tests, Hypotheses 5-9 were accepted and Hypothesis 4 was rejected. The results of all the tests of hypotheses were summarized in Table 22 in Chapter III.

From the discussion of the reliability and validity of the two measuring instruments, it was concluded that for the group of subjects as a whole, and for the second grade, the tests were adequately reliable instruments; reliability of both tests was considerably lower for the fourth grade. The tests seemed to be constructed according to valid, logical constructs based on learning theory and research on the relationship between oral language ability and reading ability.

Limitations

These limitations should be considered when evaluating the conclusions which follow:

1. The sample of oral language patterns of second and fourth graders that comprised the two reading comprehension tests devised by the investigator was not exhaustive. The selection of other frequent and infrequent patterns could conceivably yield different results.

2. As the sample used in the study was not randomly selected from the elementary school population of Madison, results should not be generalized beyond the original sample.

3. Different sets of multiple choice pictures were used in each test. The effect of this difference on the comprehension scores is not known.

4. The problem of comparing results from different grade levels was made apparent by the need to select one of two ways of constructing the tests: (1) completely different tests for each grade level, or (2) the same materials for both grades. As discussed in Chapter II, each alternative involved certain difficulties in interpreting the results. The decision to use the same set of materials for all subjects and thus control the same variables across grades may have resulted in more significant differences between grade levels than actually existed.

5. The reading comprehension tests (Test A and Test B) were considered reliable group measures for the subjects as a whole and for the second graders; the tests were considerably less reliable for the fourth grade subjects.

6. In both Test A and Test B, oral language patterns with T markers (a coordinating conjunction) were joined to form one sentence (e.g., In Test A, 1 2 4 +T and T 1 2 4 = 1 2 4 + 1 2 4). Strickland described individual patterns, not how they were joined; thus it is not known how closely the resulting combinations resembled children's use of these patterns.

Conclusions

The following conclusions appear to be warranted within the limitations of the study:

1. A significant number of second and fourth graders comprehend material written with frequent oral language patterns better than material written with infrequent oral language patterns.

2. Fourth graders comprehend material written with frequent and infrequent oral language patterns significantly better than second graders.

3. Second- and fourth-grade girls comprehend material written with frequent and infrequent oral language patterns better than boys in the same grades, but in general this difference does not reach an acceptable level of significance ($\alpha = .05$).

4. Easily administered, reliable, and valid instruments can be constructed from a relatively small sample of children's oral language patterns to determine whether or not certain patterns are more difficult for children to comprehend than others.

Implications

Implications of the findings for research in the following areas are discussed: beginning reading instruction, teacher training, classroom materials, and measurement. More specific research problems that extend from the present study are also outlined.

Implications for Beginning Reading Instruction

A developmental pattern was suggested by the finding that the difference between reading comprehension of frequent and infrequent oral patterns seemed to be greater for second graders than fourth graders. With even younger readers, therefore, the distinction between these divergent language patterns could be expected to be greater. For beginning readers, it seems logical and in keeping with linguistic knowledge to use children's patterns of language structure in written material to facilitate learning the concept that spoken and written language are related. Structuring material more like the way children speak should ease children into this new language skill more efficiently and easily than the use of material which focuses on vocabulary control to the exclusion of the kinds of language structures in which the vocabulary is used.

Proponents of language-experience approaches to teaching reading have used children's dictated speech for beginning instruction, but the concept of familiar and unfamiliar language patterns has not been crucial or even evident in these programs. Typically, as soon as the child "gets" the idea of reading, he is moved into trade books or other prepared materials without further concern for his ability to deal with specific types of sentences. In contrast to this practice, the findings of the present study suggest that children would benefit from some kind of control over sentence patterns until they are readily able to untangle word relationships in any number of infrequent patterns.

Implications for Teacher Training

The application of linguistics to the teaching of reading has been a relatively new process beginning, perhaps, with Bloomfield and Barnhart's publication Let's Read (1961). Findings of the present study support linguists who suggest that reading has something to do with the ability to comprehend the large units of language patterns as well as the small units of individual words (Lefevre, 1964). Linguists point out that traditional reading methodology is not rigorously concerned with knowledge about our language and its structure (Carroll, 1964). Extension of their principles with reference to the findings of the present study would indicate that teachers would be more effective in helping their children learn to read if they were aware of the following:

1. There is evidence that growth in oral language facility is related to growth in general reading achievement. The findings of the present study do not serve to clarify this position beyond indicating that certain oral language patterns are easier to comprehend in written material and that more experienced readers (fourth graders) do not differentiate their reading comprehension of these patterns as much as less experienced readers (second graders). The implication is that teachers should provide many opportunities for oral language growth, the assumption being that encountering patterns of structure in written material may be easier if these patterns are already in a child's repertoire.

2. Teachers should be familiar with some of the content of linguistics. They should understand the characteristics of our language—its patterns, its regularities, and its irregularities—and how these affect the relationship between spoken and written language.

3. By really listening to children, teachers could detect two possible sources of difficulty in reading comprehension. First, they could determine to what extent a child's dialect deviated from the dialect he was asked to read; major discrepancies could be expected to interfere with his understanding of the relationship between spoken and written language. Second, teachers could hear whether a child's oral reading reflected his understanding of whole groups of words—intonation patterns—rather than single words strung together.

Implications for Classroom Materials

When maximum comprehension is the goal, writers of materials for elementary school children should consider the use of language patterns that children find easier to comprehend than others. Strickland's list (1962) of frequent and infrequent oral language patterns, according to the findings of the present study, could provide a general guide to those patterns that should be relatively easy or difficult to comprehend. The guide would be modified, of course, and tempered by the specific nature of the materials, the grade level of the children being considered, and the intent of the material. The findings of this study imply that control over vocabulary is not the only logical and desirable control when comprehension of language structures is essential.

Implications for Measurement of Language-Related Abilities

From the instruments constructed for this study, it is clear that tests can be devised to measure the relative difficulty of select oral language patterns in written materials for elementary school children. Several interesting and worthwhile possibilities are consequently available for a variety of situations:

1. Since differences between reading comprehension of frequent and infrequent oral language patterns can be measured, a series of tests could be constructed to determine specific language patterns that are more or less difficult than others. Establishing a hierarchy of difficulty could be helpful in the area of readability as well as in the area of developmental reading instruction.

2. In the present study it was felt that determining reading comprehension of individual, unrelated sentences was a practical and relatively

uncomplicated procedure. By this same method, a pool of unrelated test items could be accumulated after pilot testing with a variety of grade levels and types of children. These items could be grouped in several ways, depending on the intent of the measuring instrument. For example, by using a sufficiently large number of items that dealt with the same language pattern, it could be determined whether or not an individual was having difficulty with the pattern. Or, by including a mixture of items labeled "frequent" and "infrequent," a survey could be conducted of an individual's ability to handle divergent language patterns. Construction of such tests could be helpful in determining where non-English-speaking children on the one hand, and speakers of non-standard dialects on the other were having difficulty in "straightening out" the relationships of words in English sentences.

Implications for Research

During the preparation and execution of this investigation, several research problems emerged that seem relevant and feasible in light of the findings of this study:

1. An experimental study is needed to determine if the immediate use of materials structured more like the way children speak is more beneficial to beginning readers than materials which do not control sentence patterns in this manner. The study should be longitudinal in nature, as long-term effects of such an innovative program would be important in terms of attitudes as well as reading skills.

2. Exploratory studies like this one and Ruddell's (1963) have not dealt with the reason for a gap between children's ability to speak complex structures and their ability to read these same structures. Though not a direct outgrowth from the findings of this study, research that would

illuminate the cognitive process involved in comprehending language structures is needed. With an increased understanding of the interdependence of speaking, reading, and writing, teaching procedures could become better suited to helping children grow successfully in each of these areas; methods of assessing progress could become more relevant and accurate.

3. More information is needed about the ability of children at various levels of achievement to comprehend select oral language patterns in written material. Discovering the presence of distinctly different levels of performance among good and poor readers at different grade levels would help to clarify the general findings of the present study.

4. Without the time limitation of the present study, parallel forms of tests similar to Test A and Test B could be developed for different grade levels. These tests would provide a yardstick against which a child's present level of performance could be compared to (1) his achievement the preceding year (on a test constructed for that grade level), or (2) to the achievement of other children within the same grade. Construction of a series of parallel tests would be a first step in defining "typical" ability to comprehend divergent language patterns at various levels of development in the elementary school.

5. Without sacrificing the controls imposed in Test A and Test B, it should be possible to select frequent and infrequent oral language patterns so that the same multiple choice pictures could be used to measure reading comprehension of both kinds of patterns. The use of identical pictures from test to test would provide a tighter control over the instruments.

APPENDIX A

The Twenty-five Most Frequent Oral Language Patterns of Second and Fourth Graders as Identified by Strickland (1962)

GRADE TWO ORAL LANGUAGE PATTERNS

Rank Order	Pattern	Frequency
1	1 2 4	309*
2	T 1 2 4	264*
3	1 2 4 + T	189*
4	1 2b 5	130*
5	T 1 2b 5	107
6	T 1 2 4 + T	74*
7	M_3 1 2 4	71
8	1 2b 5 + T	57
9	M_3 1 2 4 + T	52
10	1 2	48
11	1 2 M_1 + T	43
12	1 2 + T	42
13	T 1 2 M_1	38
14	1 2 4 T	37
14	T 1 2 4 M_1	37
15	T 1 2	36
16	1 2 M_1	34
17	T 1 2 M_2	32**
17	1 2 M_2	32**
18	1 2 4 M_1	31
19	T 1 2 4 M_2	29**
20	M_3 1 2 M_1 + T	28**
21	M_3 1 2 + T	27
22	T 1 2b 5 + T	25
23	T M_3 1 2 4	23
23	1 2 3 4	23**
24	T 1 2 3 4	22
24	1 2 4 M_2	22
25	1 2 4 M_1 + T	21**

* Patterns used in Test A (frequent oral language patterns)

** Patterns used in Test B (infrequent oral language patterns)

GRADE FOUR ORAL LANGUAGE PATTERNS

Rank Order	Pattern	Frequency
1	1 2 4	202*.
2	T 1 2 4	156*.
3	1 2 4 + T	123*.
4	1 2b 5	108*.
5	T 1 2 4 + T	84*
6	1 2 4 T	64
7	T 1 2b 5	63.
8	1 2 M_1 + T	51.
9	T 1 2 M_1	41
9	T 1 2 4 T	41
10	1 2 M_1	37.
11	T M_3 1 2 4 + T	33
12	1 2b 5 + T	31
13	T 1 2	29
13	M_3 1 2 4	29.
14	M_3 1 2 4 + T	28
14	1 2 4 M_1	28.
15	1 2 4 M_2	27.
16	T M_3 1 2 4	26.
17	T 1 2 4 M_1 + T	25
17	T 1 2 M_2	25**
17	M_3 1 2 M1	25.
18	1 2	23
18	T 1 2 4 M_1	23
18	T 1 2 + T	23
19	1 2 M_2	22**
19	T 1 2b 5 + T	22
20	1 2 4 M_3	21
20	T 1 2 4 M_2	21**
21	M_3 1 2 M_1 + T	19**

(Continued)

GRADE FOUR ORAL LANGUAGE PATTERNS (Continued)

Rank Order	Pattern	Frequency
21	T 2 4	19
22	1 2 M_1 T	18
23	1 2 T	17
23	1 2 M_2 + T	17
23	1 2 4 M_1 + T	17**
24	1 2 3 4	15**..
24	T M_3 1 2 M_1	15
25	T 1 2 4	14

* Patterns used in Test A (frequent oral language patterns)

** Patterns used in Test B (infrequent oral language patterns)

. Patterns used in Ruddell's high frequency passages (1963)

.. Patterns used in Ruddell's low frequency passages (1963)

APPENDIX B

The Measuring Instruments:
Directions and Reading Comprehension Tests A and B

Appendix B is not printed herein. See original manuscript in the University of Wisconsin Memorial Library.

APPENDIX C

Language Patterns in the Sentences of Test A and Test B

Appendix C is not printed herein. See original manuscript in the University of Wisconsin Memorial Library.

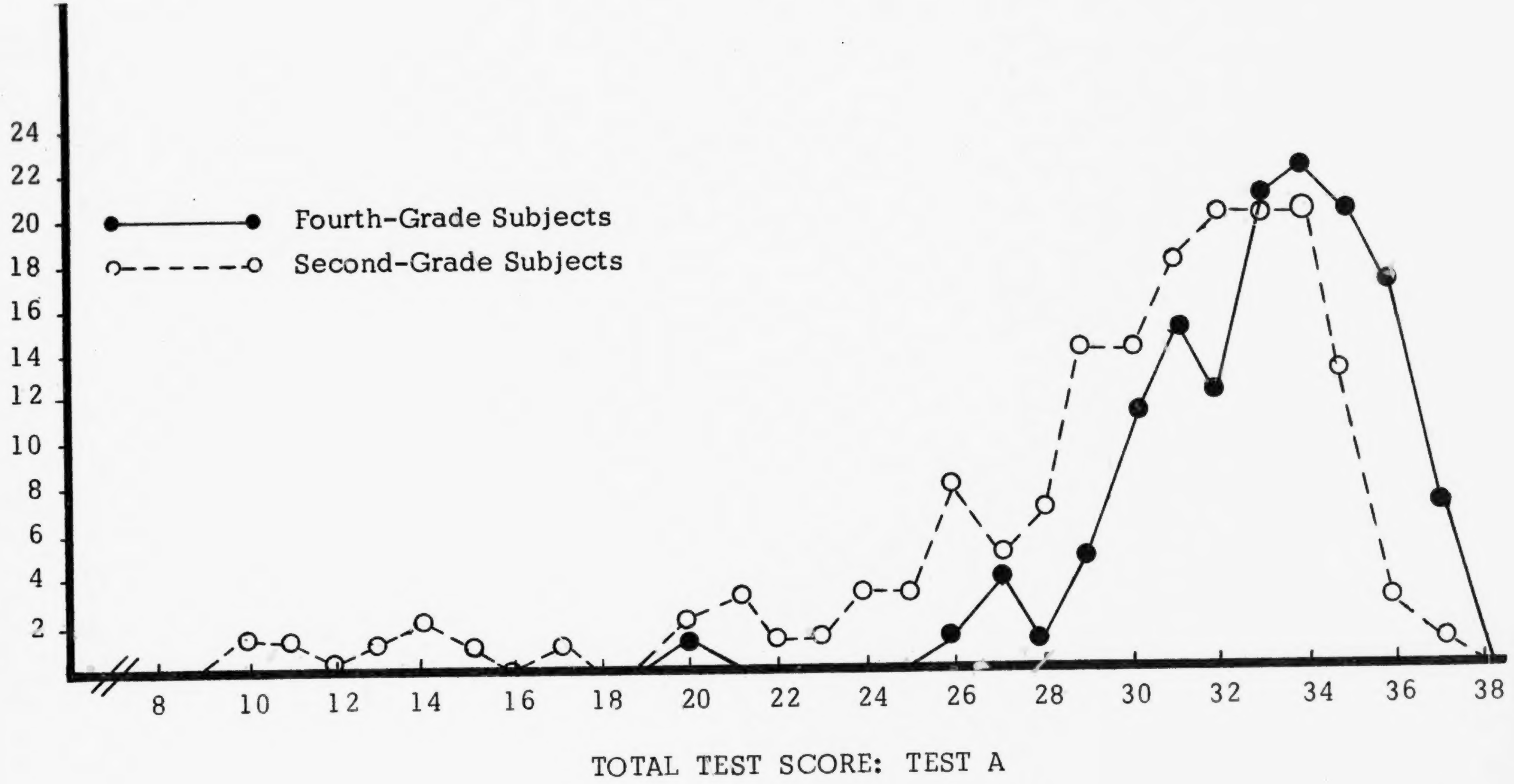
APPENDIX D

Reading Comprehension Test Scores for All Subjects

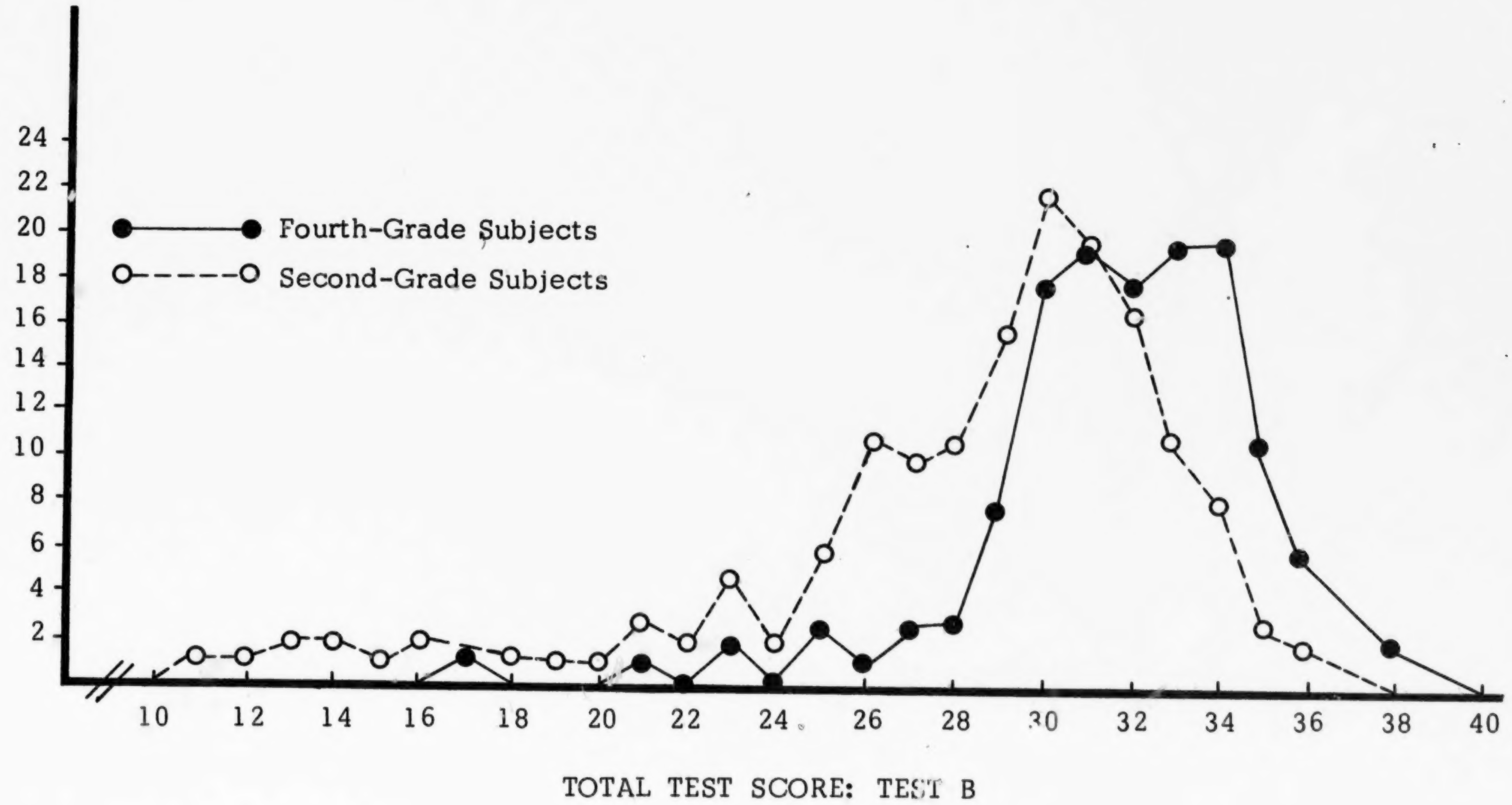
Appendix D is not printed herein. See original manuscript in the University of Wisconsin Memorial Library.

APPENDIX E

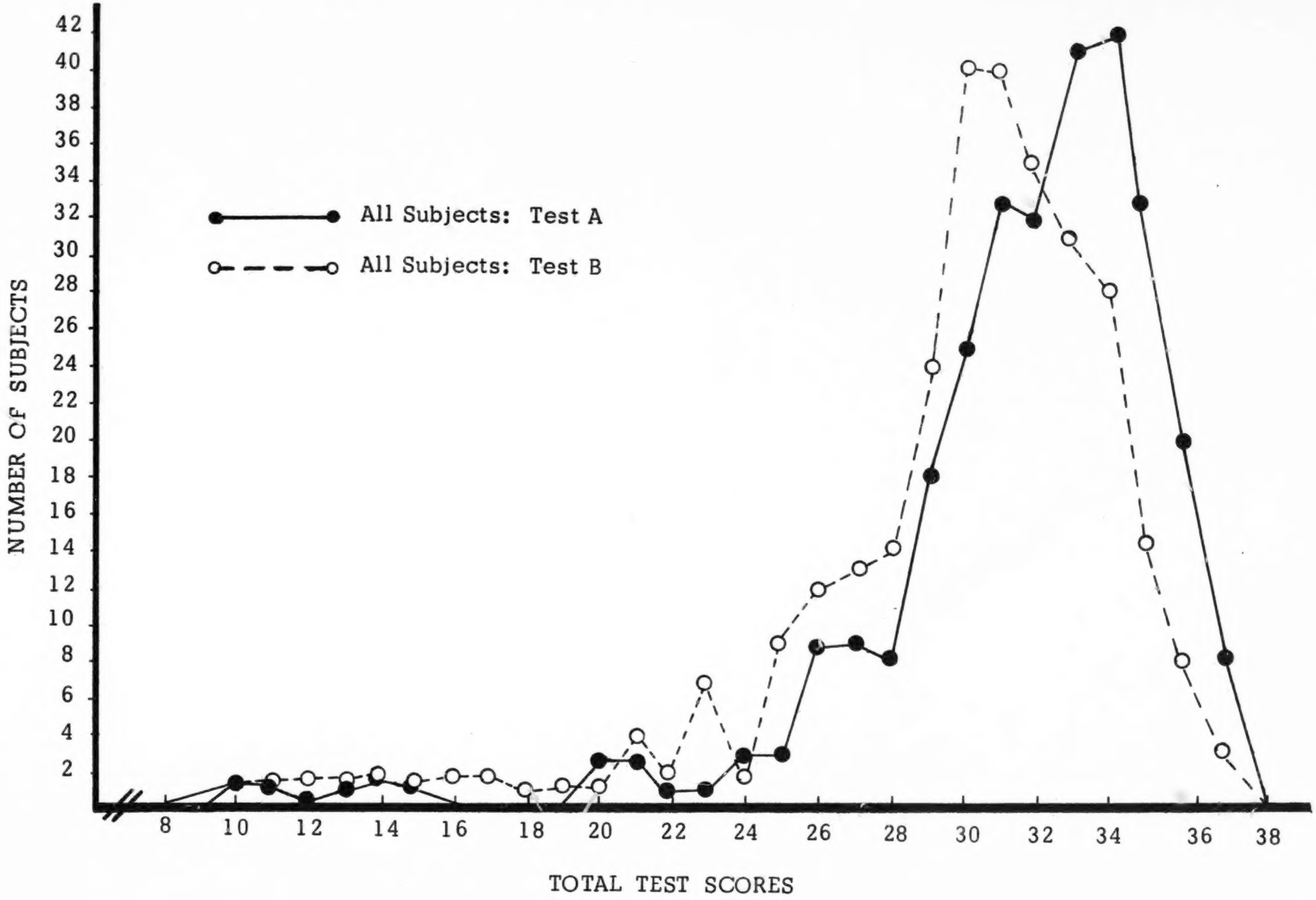
Frequency Distributions of Scores on Test A and Test B Within and Across Grades Two and Four



Frequency Distributions of Total Scores on Test A for Second and Fourth Grades



Frequency Distributions of Total Scores on Test B for Second and Fourth Grades



Frequency Distribution of Total Scores for All Subjects on Test A and Test B

APPENDIX F

Item Analysis of Test A and Test B for Grades Two and Four

Appendix F is not printed herein. See original manuscript in the University of Wisconsin Memorial Library.

BIBLIOGRAPHY

Amsden, Constance E. "Oral Language and Printed-Word Reading." Claremont College Reading Conference Twenty-ninth Yearbook. Claremont, California: Claremont College Curriculum Laboratory, 1964, pp. 90-3.

Anastasi, Anne. Psychological Testing. New York: The Macmillan Company, 1961.

Baker, Frank B. and T. J. Martin. FORTAP, A Fortran Test Analysis Package. The University of Wisconsin: Wisconsin Research and Development Center for Cognitive Learning, March 1, 1968.

Barrett, Thomas C. The Relationship Between Selected Reading Readiness Measures of Visual Discrimination and First Grade Reading Achievement. Unpublished doctoral dissertation. University of Minnesota, 1962.

Barrett, Thomas C. and Wayne Otto. "Elementary Pupil's Ability to Conceptualize the Main Idea in Reading." A paper presented at the American Educational Research Association meeting, Chicago, February, 1968.

Betts, Emmett Albert. "Readability: Its Application to the Elementary School." Journal of Educational Research. 42:6, February, 1949, pp. 439-59.

Bloom, Benjamin. "Test Reliability for What?" Readings in Educational and Psychological Measurement. (ed. by Clinton Chase and H. Glenn Ludlow), New York: Houghton Mifflin Company, 1966, pp. 126-34.

Bloomfield, Leonard and Clarence Barnhart. Let's Read. Detroit: Wayne State University Press, 1961.

Burrows, Alvina. "Research Critiques." Elementary English. 41, May, 1964, pp. 532-5.

Carroll, John B. "The Analysis of Reading Instruction: Perspectives from Psychology and Linguistics," Theories of Learning and Instruction (ed. by Ernest R. Hilgard), Chicago, Illinois: The National Society for the Study of Education, 1964, pp. 336-53.

______. "Communication Theory, Linguistics, and Psycholinguistics," Review of Educational Research. 28:2, April, 1958.

______. "Language Development," Encyclopedia of Educational Research. (ed. by Chester W. Harris), New York: The Macmillan Company, 1960, pp. 283-92.

______. "Linguistics and the Psychology of Language," Review of Educational Research. 34:2, April, 1964.

Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Brothers, Publishers, 1960.

Davis, O. L., Jr. and Joan G. Seifert. "Some Linguistic Features of Five Literature Book for Children." Elementary English. December, 1967.

Ebel, Robert. In Testing Problems in Perspective. (ed. by Anne Anastasi), American Council on Education, 1966, pp. 115-30.

Fries, Charles. Linguistics and Reading. New York: Holt, Rinehart and Brothers, Publishers, 1963.

Goodman, Kenneth. Study of Children's Behavior While Reading Orally. U. S. Department of Health, Education, and Welfare, Office of Education, Project No. S 425, Detroit, Michigan: Wayne State University, 1968.

Gray, William S. "Progress in the Study of Readability." The Elementary School Journal. May, 1947, pp. 491-99.

Guha, Susanta. "Multiple Regression Analysis REGANI," The University of Wisconsin Computing Center, Users Manual, Vol. 6 - Preliminary, Statistical Programs for the 1604 Computer. Nov. 1966, Section 3.2.

Hays, William L. Statistics for Psychologists. New York: Holt, Rinehart and Winston, 1963.

Hocker, Mary Elsa. Reading Materials for Children Based on Their Language Patterns of Syntax, Vocabulary, and Interests. Unpublished Master's thesis, University of Arizona, 1963.

Hodge, Robert, Paul Siegel, and Peter Rossi. "Occupational Prestige in in the United States, 1925 to 1963," American Journal of Sociology, 70:286-302, November, 1964.

Hunt, Kellogg W. "Recent Measures in Language Development." A paper presented at the National Conference on Research in English, Chicago, February, 1966.

______. Grammatical Structures Written at Three Grade Levels. U. S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 1998. Tallahasee, Florida: Florida State University, 1965.

Kean, John. The Linguistic Structure of Second and Fifth Grade Teachers' Oral Classroom Language. U. S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. S-331, Madison, Wisconsin: University of Wisconsin, 1967.

Kelley, Truman L., and Richard Madden, Eric Gardner, and Herbert Rudman. Stanford Achievement Test: Intermediate I and Primary II, Form Y. New York: Harcourt, Brace and World, Inc., 1964.

Kjeldergaard, Paul M. "The Psychology of Language," Review of Educational Research. 31:2, April, 1961.

Lefevre, Carl. Linguistics and the Teaching of Reading. New York: McGraw-Hill Book Company, 1964.

Lindgren, Henry Clay. Educational Psychology in the Classroom. New York: John Wiley and Sons, Inc., 1962.

Lindvall, C. M. Measuring Pupil Achievement and Aptitude. New York: Harcourt, Brace and World, Inc., 1967.

Loban, Walter. The Language of Elementary School Children. National Council of Teachers of English, Research Report No. 1. Champaign, Illinois: the Council, 1963.

Mayer, Edgar. "Research Critiques." Elementary English. 41, May, 1964, pp. 535-36, + 541.

McCarthy, Dorothea. "Language Development in Children." A Manual of Child Psychology. (ed. by Leonard Charmichael) New York: John Wiley and Sons, Inc., 1954.

Mortenson, William Paul. Differences in Performance of Beginning First Grade Pupils on Selected Pre-Reading Tasks According to Socio-Economic Status and Sex. Unpublished doctoral dissertation, University of Wisconsin, Madison, 1966.

O'Donnell, Roy and William Griffin, and Raymond Norris. Syntax of Kindergarten and Elementary School Children: A Transformational Analysis. National Council of Teachers of English, Research Report No. 8. Champaign, Illinois: the Council, 1967.

Reiss, Albert J. Occupations and Social Status. New York: The Free Press of Glencoe, Inc., 1961.

Riling, Mildred. Oral and Written Language of Children in Grades 4 and 6 Compared with the Language of Their Textbooks. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 2410. Durant, Oklahoma: Southeastern State College, 1965.

Ruddell, Robert Byron. An Investigation of the Effect of the Similarity of Oral and Written Patterns of Language Structure on Reading Comprehension. Unpublished doctoral dissertation, Indiana University: 1963.

______. The Effect of Four Programs of Reading Instruction with Varying Emphasis on the Regularity of Grapheme-Phoneme Correspondences and the Relation of Language Structure to Meaning on Achievement in First Grade Reading. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 2699. Berkeley, California: University of California, 1965.

Stevens, Martin. "Intonation in the Teaching of Reading." Elementary English. 42:231-37, March, 1965.

Stone, C. R. "Measuring Difficulty of Primary Reading Material: A Constructive Criticism of Spache's Measure." Elementary School Journal. 57:410-13, March, 1953.

Strickland, Ruth G. "The Language of Elementary School Children: Its Relationship to the Language of Reading Textbooks and the Quality of Reading of Selected Children," Indiana University School of Education Bulletin. 38: 1-131, July, 1962.

Tate, Merle. Statistics in Education. New York: The Macmillan Company, 1955.

Tatham, Susan M. "The Effect of Material Written with Certain Oral Language Patterns on Reading Comprehension." A paper presented at the American Educational Research Association meeting, New York City, February, 1967.

______. "Using Multiple Choice Questions to Measure the Effect on Comprehension of Material Written with Select Oral Language Patterns." A paper presented at the American Educational Research Association meeting, Chicago, February, 1968.

Templin, Mildred. "Certain Language Skills in Children: Their Development and Interrelationships." Institute of Child Welfare Monograph Series, No. 26. Minneapolis: University of Minnesota Press, 1957.

Vygotsky, Lev Semenovich. Thought and Language. Cambridge, Mass.: Massachusetts Institute of Technology Press, 1962.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1962.